

JTRP

Joint
Transportation
Research
Program

FHWA/IN/JTRP-97/12

Final Report

PART I

**MODELING AND OPTIMIZATION OF
THE INDIANA LANE MERGE CONTROL
SYSTEM ON APPROACHES TO
FREEWAY WORK ZONES**

A. Tarko
S. Kanipakapatnam
J. Wasson

May 1998

Indiana
Department
of Transportation

Purdue
University

Final Report

FHWA/IN/JTRP-97/12

Part I

**Modeling and Optimization of the Indiana Lane Merge Control System on Approaches to
Freeway Work Zones**

by

Andrzej P. Tarko
Professor

Sreenivasulu R. Kanipakapatnam
Research Assistant

Jason S. Wasson
Research Assistants

**Transportation and Infrastructure Systems
Department of Civil Engineering
Purdue University**

**Joint Transportation Research Program
Project No: C-36-75F
File No: 8-9-6**

**Prepared in Cooperation with the
Indiana Department of Transportation and
the U.S. Department of Transportation
Federal Highway Administration**

The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration and the Indiana Department of Transportation. This report does not constitute a standard, specification or regulation.

**Purdue University
West Lafayette, IN 47907
May 1998**

Digitized by the Internet Archive
in 2011 with funding from
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

1. Report No. FHWA/IN/JTRP-97/12	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Modeling and Optimization of the Indiana Lane Merge Control System on Approaches to Freeway Work Zones		5. Report Date May 1998	
		6. Performing Organization Code	
7. Author(s) A. Tarko, S. Kanipakapatnam, and J. Wasson		8. Performing Organization Report No. FHWA/IN/JTRP-97/12	
9. Performing Organization Name and Address Joint Transportation Research Program Purdue University West Lafayette, Indiana 47907-1284		10. Work Unit No.	
		11. Contract or Grant No. SPR-2127	
12. Sponsoring Agency Name and Address Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204		13. Type of Report and Period Covered Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes Prepared in cooperation with the Indiana Department of Highways and Federal Highway Administration.			
16. Abstract <p>Severe traffic turbulence on entry sections of freeway work zones increases the delays and risk of crash. A new Indiana Department of Transportation system called Indiana Merge Lane System (IMLS) creates a dynamic no passing zone on the approach to the freeway work zone through the sequence of DO NOT PASS signs. The system is thought to encourage drivers to switch lanes well upstream of the discontinuous lane taper where the merging maneuver is safer and less intrusive. The IMLS is expected to impact drivers' behavior, their perception of the traffic conditions, and traffic safety.</p> <p>This research is focused on: (1) drivers' compliance with the system, (2) delays and travel times on approaches to work zones, (3) optimal configuration of the system, and (4) warrants for the system's use. The simulation and field studies indicate a significant reduction in the number of merging maneuvers near work zones after the IMLS is applied. Also, the travel time on continuous lanes is reduced. The increased fairness of the system improves the perception of the traffic conditions among the majority of drivers. A slight reduction in the capacity of the merge point is the second finding of the field observations. This finding should be confirmed through long-term measurements of capacity during regular use of the IMLS units.</p> <p>The final report is divided into two parts. Part I presents the performed research, including the simulation model development and simulation experiments. Part II contains the system description, guidelines for its use, and rules for its setting. The system description includes presentation of the concept and the system components. The guidelines for the system use provide the traffic conditions where the system is expected to provide benefit. Finally, the manual gives a set of simple rules useful in setting all the system parameters to achieve the maximum reduction in the travel time in the continuous lane.</p>			
17. Key Words freeway work zone, traffic control, lane merge, simulation		18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Virginia, 22161	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 128	22. Price

TABLE OF CONTENTS

TABLE OF CONTENTS.....	I
LIST OF FIGURES.....	IV
LIST OF TABLES	VI
IMPLEMENTATION REPORT	VII
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 SIMULATION MODEL.....	4
2.1 BACKGROUND.....	4
2.1.1 Problem Definition and Objectives.....	5
2.1.2 Approach	5
2.2 LITERATURE REVIEW	6
2.2.1 Major Approaches of Modeling.....	6
2.2.2 Macroscopic Modeling	6
2.2.2.1 Simple Continuum Models	8
2.2.2.2 Higher-order Continuum Models.....	9
2.3 EXTENSIONS OF THE CONTINUUM MODEL	11
2.3.1 Basic Relationship	12
2.3.2 Basic Continuity Equation	12
2.3.3 Momentum Equation	13
2.3.4 Lane Change Rate.....	14
2.3.4.1 Proportion of Drivers Changing Lanes	15
2.3.4.2 Time for Lane Change Maneuver	18
2.3.5 Desirable Speed	19
2.4 MODEL DISCRETIZATION	21
2.4.1 Background.....	21
2.4.2 Basic Relationship	22
2.4.3 Basic Continuity Equation	24

2.4.4	Momentum Equation	24
2.4.5	Boundary Conditions	25
2.4.6	Initial Conditions	25
2.4.7	Solution Approach	26
2.5	DATA COLLECTION.....	28
2.5.1	Site Selection	28
2.5.2	Techniques Applied in Data Collection	29
2.5.2.1	Nu-Metrics Detectors.....	29
2.5.2.2	Distance Measuring Vehicle	30
2.5.3	Data Processing	30
2.5.4	Data Collection with Control in Place	31
2.5.5	Discussion of the Results of Field Data Collection.....	35
2.6	CALIBRATION OF THE MODEL.....	40
2.6.1	Calibration Process	40
2.6.2	First Stage of Calibration	41
2.6.3	Second Stage of Calibration	42
2.6.3.1	Method and Algorithm.....	45
2.6.3.2	Calibration Results.....	46
2.6.4	Third Stage of Calibration	54
2.7	EVALUATION OF THE MODEL	61
2.7.1	Method.....	61
2.7.2	Evaluation Results	62
2.8	SUMMARY AND CONCLUSIONS	66
2.9	LIST OF REFERENCES	68
CHAPTER 3	SIMULATION USER'S MANUAL.....	70
3.1	INTRODUCTION	70
3.2	INPUT FILES	70
3.2.1	D_Chars.dat	71
3.2.2	F_Chars.dat.....	73
3.2.3	C_Chars.dat	74
3.2.4	O_Chars.dat	76

3.2.5 lane1.inp & lane2.inp.....	76
3.3 OUTPUT FILES	77
3.3.1 detail.out	77
3.3.2 overall.out.....	77
3.3.3 C_In_tme.out	78
3.3.4 report.out	78
 CHAPTER 4 SIMULATION EXPERIMENT	 79
4.1 OBJECTIVES OF THE SIMULATION	79
4.2 SIMULATED CASES	79
4.3 MEASURE OF EFFECTIVENESS	80
4.4 OPTIMIZATION OF INDIVIDUAL PARAMETERS.....	82
4.5 VARIABLES' CONSTRAINTS	85
4.6 SIMULATION RESULTS	85
4.7 PROPOSED RULES OF SYSTEM SETTING	88

LIST OF FIGURES

Figure 1.1 Concept of the Indiana Lane Merge System on Approach to Freeway Work Zone.....	3
Figure 2.3.1 Decision tree for lanes change.....	16
Figure 2.4.1 Sample Layout of a Two-lane Freeway with One Lane Dropped.....	23
Figure 2.4.2 A Typical Section of a Freeway Lane.....	23
Figure 2.4.3 Flowchart for Calculating Flows, Speeds, and Densities.....	27
Figure 2.5.1 Countcard Detector.....	32
Figure 2.5.2 Hi-Star Detector	32
Figure 2.5.3 Layout of Detectors for Data Collection without Control	33
Figure 2.5.4 Layout of Detectors for Data Collection with Control.....	34
Figure 2.5.5 Variation of Total Inflow on I-69 Southbound Approach	36
Figure 2.5.6 Variation of Traffic Distribution between Lanes on I-69 Southbound Approach.....	36
Figure 2.5.7 Density Variation 1.25 km to Lane Drop (Detector Location 2)	37
Figure 2.5.8 Density Variation with Time at 0.05 km to Lane Drop (Detector Location 3).....	37
Figure 2.5.9 Flow Variation with Time at 1.25 km to Lane Drop (Detector Location 2)	38
Figure 2.5.10 Flow Variation with Time at 0.05 km to Lane Drop (Detector Location 3)	38
Figure 2.5.11 Speed Variation with Time at 1.25 km to Lane Drop (Detector Location 2)	39
Figure 2.5.12 Speed Variation with Time at 0.05 km to Lane Drop (Detector Location 3)	39
Figure 2.6.1 Scatter Plot of Density and Speed for I-69 Southbound Freeway Approach.....	43
Figure 2.6.2 Scatter Plot of Density and Speed for I-69 Northbound Freeway Approach.....	43
Figure 2.6.3 Proposed Relationship between Density and Speed (I-69 Northbound Freeway Data).....	44
Figure 2.6.4 Flow Variation with Time at 1.25 km to Lane Drop in Continuous Lane	49
Figure 2.6.5 Speed Variation with Time at 1.25 km to Lane Drop in Continuous Lane.....	49
Figure 2.6.6 Density Variation with Time at 1.25 km to Lane Drop in Continuous Lane	49
Figure 2.6.7 Flow Variation With Time at 1.25 km to Lane Drop in Dropped Lane.....	50
Figure 2.6.8 Speed Variation With Time at 1.25 km to Lane Drop in Dropped Lane	50
Figure 2.6.9 Density Variation With Time at 1.25 km to Lane Drop in Dropped Lane	50
Figure 2.6.10 Flow Variation with Time at 0.05 km to Lane Drop in Continuous Lane	51
Figure 2.6.11 Flow Variation with Time at 0.05 km to Lane Drop in Continuous Lane	51
Figure 2.6.12 Density Variation with Time at 0.05 km to Lane Drop in Continuous Lane	51
Figure 2.6.13 Flow Variation with Time at 0.05 km to Lane Drop in Dropped Lane.....	52
Figure 2.6.14 :Speed Variation with Time at 0.05 km to Lane Drop in Dropped Lane.....	52
Figure 2.6.15 Density Variation with Time at 0.05 km to Lane Drop in Dropped Lane	52
Figure 2.6.16 Flow Variation with Time at -0.80 km to Lane Drop in Continuous Lane	53
Figure 2.6.17 Speed Variation with Time at -0.80 km to Lane Drop in Continuous Lane	53

Figure 2.6.18 Density Variation with Time at -0.80 miles to Lane Drop in Continuous Lane	53
Figure 2.6.19 Flow Variation with Time at 1.25 km to Lane Drop in Continuous Lane	56
Figure 2.6.20 Speed Variation With Time at 1.25 km to Lane Drop in Continuous Lane.....	56
Figure 2.6.21 Density Variation With Time at 1.25 km to Lane Drop in Continuous Lane	56
Figure 2.6.22 Flow Variation With Time at 1.25 km to Lane Drop in Dropped Lane.....	57
Figure 2.6.23 Speed Variation With Time at 1.25 km to Lane Drop in Dropped Lane	57
Figure 2.6.24 Density Variation With Time at 1.25 km to Lane Drop in Dropped Lane	57
Figure 2.6.25 Flow Variation With Time at 0.05 km to Lane Drop in Continuous Lane	58
Figure 2.6.26 Speed Variation With Time at 0.05 km to Lane Drop in Continuous Lane.....	58
Figure 2.6.27 Density Variation With Time at 0.05 km to Lane Drop in Continuous Lane	58
Figure 2.6.28 Flow Variation With Time at 0.05 km to Lane Drop in Dropped Lane.....	59
Figure 2.6.29 Speed Variation With Time at 0.05 km to Lane Drop in Dropped Lane.....	59
Figure 2.6.30 Density Variation With Time at 0.05 km to Lane Drop in Dropped Lane	59
Figure 2.6.31 Flow Variation With Time at -0.80 km to Lane Drop in Continuous Lane	60
Figure 2.6.32 Speed Variation With Time at -0.80 km to Lane Drop in Continuous Lane	60
Figure 2.6.33 Density Variation With Time at -0.80 km to Lane Drop in Continuous Lane	60
Figure 2.7.1 Flow Variation at Input Section Both in Continuous and Dropped Lanes.....	63
Figure 2.7.2 Density Variation With Time at 0.85 miles to lane drop in Continuous Lane	64
Figure 2.7.3 Density Variation With Time at 0.85 miles to lane drop in Dropped Lane	64
Figure 2.7.4 Density Variation With Time at 0.06 miles to lane drop in Continuous Lane	65
Figure 2.7.5 Density Variation With Time at 0.06 miles to lane drop in Dropped Lane	65
Figure 2.7.6 Density Variation With Time at -0.91 miles to lane drop in Continuous Lane.....	66
Figure 4.1 Illustration of Traffic Demand and Work Zone Capacity	80
Figure 4.2 Results of the Simulation Experiments for $V = 1600$ veh/h	93
Figure 4.3 Results of the Simulation Experiments for $V = 1700$ veh/h	93
Figure 4.4 Results of the Simulation Experiments for $V = 1800$ veh/h	94
Figure 4.5 Maximum Congested Segment versus Deployment Area.....	97

LIST OF TABLES

Table 2.6.1 Traffic Parameters' Constraints (distance is in meters and time is in seconds)	47
Table 2.6.2 Calibrated Traffic Parameters (distance is in meters and time is seconds)	48
Table 2.6.3 Control Parameters' Constraints.....	55
Table 2.6.4 Calibrated Control Parameters.....	55
Table 4.1 Simulation Experiment (no control)	86
Table 4.2 Simulation Experiments to Optimize the System (Threshold Occupancy = 30%, Min. Activation = 5 min.).....	91
Table 4.3 Effect of Spacing Uniformity (spacing of static boards = 150, 450, 750m; total distance with dynamic boards = 3000 m; number of boards = 5).....	94
Table 4.4 Effect of Threshold Occupancy (spacing of static boards = 300 m).....	95
Table 4.5 Average Traffic Densities on Approaches to Work Zones with Optimal ILMS	95
Table 4.6 Additional Simulation Experiments.....	96
Table 4.7 Calculated Maximum Congested Segments versus Optimal Deployment Areas	97

IMPLEMENTATION REPORT

The final report is divided into two parts. Part I presents the performed research including the simulation model development and simulation experiments. Part II, called “ILMS Manual,” is a direct implementation of the research findings and is for traffic engineers and highway designers who participate in planning, design, and management of work zones on rural freeways. The manual contains a description of Indiana Lane Merge System (ILMS), guidelines for its use, and rules for its setting. The system description includes presentation of the system concept and system components. The guidelines for the system outline the traffic conditions where the system is expected to provide benefit. Finally, the manual gives rules useful in setting the system parameters to achieve the maximum reduction in travel time in the continuous lane.

The INDOT representatives who will take the lead in implementing the research results will be the ITS Program Engineer in the Executive Offices and the Specialty Project Group Manager in the Design Division. The criteria for selecting candidate construction projects and the method of deployment must be identified in order to implement the research results. A standard and specifications will be developed for future application of the ILMS and, as a means of providing information about the ILMS, INDOT will prepare a short video.

CHAPTER 1 INTRODUCTION

Each year in the United States transportation agencies are faced with the challenge of managing traffic in highway work zones. The biggest issues that need to be addressed with a proper traffic management are excessive travelers' delays and risk of crashes. These problems occur mainly on the sections of roadway preceding the work zones in which one or more lanes are discontinued. These areas can typically be characterized by severe traffic turbulence, resulting in increased accident risks and long traveler delays.

The entry section of the work zone is critical for traffic smoothness and safety because of the discontinuation of one or more traffic lanes. Some drivers try to avoid dense traffic in the continuous lanes by approaching the work zone entry point in the discontinued lane up to the point where the lane change maneuver is difficult and risky. The aggressive lane changes resulting from such behavior create turbulence in the traffic stream, which negatively affects traffic performance. These negative effects include shock waves in the continuous lane, as well as the development of road rage among typically non-aggressive drivers. All of these effects create an extremely dangerous situation both at the merge point, as well as within the work zone, due to the road rage, which continues far beyond the point at which the aggressive lane change takes place.

The Indiana Department of Transportation (INDOT) has undertaken an effort to implement a novel traffic control system at entries to freeway work zones. The concept was developed by Daniel E. Shamo, P.E. and James M. Poturalski, P.E. and is believed to reduce the number of aggressive lane changes by encouraging drivers to switch lanes well upstream of the discontinuous lane taper. This allows drivers who are merging into the continuous lane to safely make the maneuver because of the increased headway between vehicles and the lower differential in speed between the two lanes. The system, which is called the Indiana Lane Merge System (ILMS) consists of a series of static and dynamic signs which create a variable no passing zone in advance of a freeway work zone (see

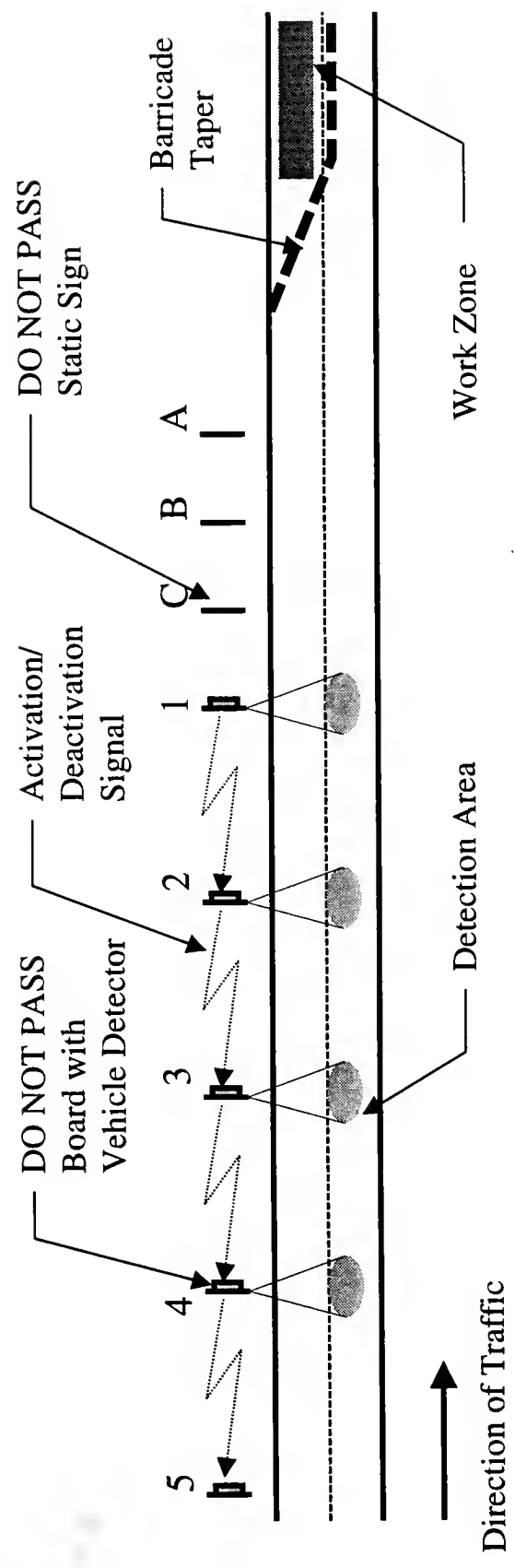
Figure 1.1). The signs are placed upstream from the taper, adjacent to the discontinuous lane. Vehicle detectors spread along the continuous lane detect traffic queues and activate dynamic signs upstream to the congested section, thereby imposing a no-passing .

According to the authors, this is a new system that has not been implemented yet. It was felt that this system could have an impact on user costs, drivers' behavior and their perception of the traffic conditions, and traffic safety. The scope of the research has addressed the lack of an automated prototype system during this project. For example, the effectiveness of the system in reducing accidents was not evaluated due to the lack of accident records. It is therefore reasonable to consider this project a first phase of a series of projects related to this system. There were four objectives of the first phase:

1. drivers' compliance with the system,
2. delays and travel times on approaches to work zones,
3. optimal configuration of the system, and
4. warrants for the system's use.

The research approach for this project consisted of several tasks. First, it was necessary to develop a computer model that could simulate the traffic characteristics with and without ILMS in operation. This model needed to be verified with field data. After the model was developed to accurately replicate the field observations, several simulation cases were performed for a variety of traffic conditions to develop general rules for optimal setting of the system. Once the system layout and settings were optimized, an evaluation of the system's effect on motorists' travel time was conducted.

This report is divided into two parts. Part I presents the research methodology, including system optimization and evaluation. Part II, the "ILMS Manual," contains guidelines for optimal setting of the system and is intended to function as a stand-alone document for users of this system. It provides the intended purpose of the system and set of rules and advises the proper geometric layout and parameter settings of the system.



Note: Board 1 is constantly on

Figure 1.1 Concept of the Indiana Lane Merge System on Approach to Freeway Work Zone

CHAPTER 2 SIMULATION MODEL

2.1 BACKGROUND

The interstate highways network -- a strategic component of the U.S. transportation system -- plays a crucial role in the nation's economy. Increasing congestion and the adverse environmental impact of the freeways has been a growing concern recently. In response, new technologies in communication and computer science, and a generation of traffic management and control systems, called Intelligent Transportation Systems (ITS), have emerged in the transportation area. This has created an opportunity to design a new generation of traffic management and control systems pertaining to freeways, such as Advanced Traveler Information Systems, Incident Management, and Freeway Work Zone Traffic Management.

The new technologies can be used to enhance traffic management and control near work zones on freeways where additional crash risks are created that sometimes cause severe delays. Entry sections of work zones, where one or more traffic lanes are discontinued, are critical for traffic smoothness and safety. Some drivers try to avoid dense traffic by approaching the work zone entry in the discontinued lane up to the point where the lane change maneuver is difficult and risky. The aggressive lane changes resulting from such behavior involve traffic conflicts and turbulence that negatively affect traffic performance. To mitigate this problem, an effort has been made recently by the Indiana Department of Transportation to implement a novel traffic control system shown in Figure 2.1. The system uses boards to display DO NOT PASS messages. The variable signs are activated by the high occupancy of the detectors placed downstream to these signs. This system is intended to reduce the number of aggressive lane changes by encouraging drivers to switch lanes earlier. The effectiveness of the new merge control and its optimal settings are unknown.

2.1.1 Problem Definition and Objectives

As mentioned in the previous section, there is a need for evaluating the effectiveness of the new merge control system applied to work zones. Design of the new control system for work zones involves decisions on the number of DO NOT PASS boards, their location, and the threshold values to activate DO NOT PASS boards. The control system can be evaluated by comparing total travel times, average travel speeds, or other synthetic measures of effectiveness. Unfortunately, there are no closed-form functions linking these measures with the traffic control settings. The system performance must be evaluated through simulation.

The aim of this study is to develop a simulation model that is suitable for the optimization of the traffic control for work zones and for studying the effectiveness of the optimal control. The simulation model must be fast and should incorporate drivers' behavior in response to various factors specific to work zones, i.e., the distance to the lane drop, opportunities for lane change, and the effect of DO NOT PASS signs.

2.1.2 Approach

First, the simulation model is developed, with the lane change behavior sensitive to such characteristics as the distance to the lane drop, presence of DO NOT PASS signs, and traffic conditions. The model has several parameters that must be calibrated using real data. In order to calibrate the model, field data are collected. The model is then calibrated and evaluated.

2.2 LITERATURE REVIEW

In this section, a brief overview of the past research in the area of traffic simulation modeling is presented. This chapter is concluded by selecting the research approach to our problem.

2.2.1 Major Approaches of Modeling

Past research in the area of freeway traffic modeling evolved with two major approaches:

- (1) microscopic models
- (2) macroscopic models

Microscopic models describe the behavior of individual vehicles as a function of traffic conditions in their environment (Herman et al., 1959; May and Keller, 1967; Gazis 1974). Microscopic models usually keep track of each vehicle in the highway system, which requires a large amount of input data. This is unfortunate because, for a simulation to work, microscopic details must be just right. Otherwise, as is well known from nonlinear system theory, microscopic details have a way of affecting the macroscopic world unpredictably (Daganzo, 1993). Moreover, since microscopic models represent each vehicle separately, the simulation is time consuming and the results are burdened with random bias. Thus, we focus on macroscopic models here, and the following section discusses the state of the art in macroscopic modeling.

2.2.2 Macroscopic Modeling

In macroscopic models, traffic is represented as a stream by using easy to observe traffic conditions routinely used in traffic engineering analyses: traffic flow, density, and speed. Macroscopic models assume that the aggregate behavior of sets of vehicles depends on

the traffic conditions in their environment. The idea of describing traffic with continuous characteristics, traditionally used to describe fluid mechanics, was first proposed by Lighthill and Whitham (1955a; 1955b). Many other researchers thereafter confirmed that macroscopic models of freeway traffic could be useful in certain applications for estimation and prediction of traffic characteristics.

Looking from an airplane at a freeway, one can visualize the line of vehicles as a stream or a continuum fluid. Since macroscopic models use this analogy, these models are sometimes referred to as continuum models. From the fluid flow analogy, the traffic stream is treated as one-dimensional compressible fluid. This leads to two basic assumptions:

- (1) Traffic is conserved, and
- (2) There is a one-to-one relationship between speed and density.

The first assumption is imposed through the continuity (conservation) equation. This principle is generally accepted and there is no controversy on its validity. The second assumption, however, has raised objections in the literature, partly because it is not always understood and its measurements are contradictory. Specifically, if speed is a function of density, it follows that drivers adjust their speed according to the density, i.e., as density increases, speed decreases. This is intuitively true, but it can theoretically lead to negative speeds or densities. In addition, it has been observed that for the same value of density, many values of speed can be observed (TRB, 1975; TRB, 1996). Evidently, the assumption requires the qualification that speed is a function of density, but only at equilibrium. Because equilibrium can rarely be observed in practice, a satisfactory speed-density relationship is hard to obtain, and it is often assumed or inferred theoretically. The simple continuum model consists of the conservation equation, and the equilibrium speed-density relationship. If these two equations are solved together with the basic relationship (flow is equal to speed times density), then we can obtain flow, speed, and density at any instant and point on a freeway. The solution of the simple continuum model involves the generation of shock waves. This problem has been solved by

considering higher-order continuum models. The higher-order models consider momentum equation that accounts for the acceleration and inertia characteristics of traffic mass and therefore can be applied to non-equilibrium flows. In spite of this improvement, the most widely known higher-order models still require equilibrium speed-density relationship. Recently, new higher-order models were proposed that remove this requirement, but they are largely untested (Lyrintzis A.S. et al, 1994).

The following Sections 2.2.2.1 and 2.2.2.2 describe simple continuum models and higher-order models.

2.2.2.1 Simple Continuum Models

As mentioned in the previous section, simple continuum models consist of conservation equation and the equation of state (speed-density relationship for equilibrium conditions). The following are the three equations that constitute simple continuum model.

(1) Fundamental relationship between the three traffic characteristics:

$$q = k \cdot v \quad (2.1)$$

(2) Basic continuity equation ensures the conservation of traffic mass:

$$\frac{\partial q}{\partial x} + \frac{\partial k}{\partial t} = g(x, t) \quad (2.2)$$

where:

$g(x, t)$ is the generation/dissipation rate in vehicle per unit time per unit length, and
 $(g(x, t) = 0$ if there is no generation/loss of vehicles through on/off ramps)

Equation 2.2 is also referred to as conservation equation. The physical interpretation of this equation is discussed in Chapter 3.

(3) Equilibrium speed-density relationship is as follow:

$$U = U_f \cdot \left(1 - \left(\frac{k}{k_j}\right)^\alpha\right)^\beta \quad (2.3)$$

where

U_f = free flow speed,
 k_j = jam density, and
 α, β = positive constants.

For instance, for $\alpha = \beta = 1$, the above equation becomes Greenshields (1934) equation of state. The relationship 2.3 can be obtained from empirical data or sometimes theoretically inferred. In reality, there is no agreement as to the specific form of the equilibrium speed-density relationship, partly because it is tedious to obtain and traffic equilibrium rarely exists.

Equation 2.2 is the state equation that can be used to determine the flow at any point of the freeway. The attractiveness of this equation is that it relates two traffic characteristics through two independent variable, time, t , and space, x . Analytical solution based on the above set of three equations is very involved and impractical for real-life applications. The set of equations are generally solved numerically (Michalopoulos et al., 1987; Michalopoulos et al., 1991). Computation of k , q , and U proceeds in a discretized space (short intervals Δx) over discretized time (short intervals Δt). A practical implementation of the simple continuum theory to freeways and intersections can be found in (Stephanopoulos and Michalopoulos, 1981; Michalopoulos, 1988; Michalopoulos et al 1991).

2.2.2.2 Higher-order Continuum Models

The higher order models take into account acceleration/deceleration and inertia characteristics of traffic mass by replacing the equilibrium speed-density relationship with a momentum equation (Whitham, 1974; Payne, 1979). This equation is obtained by

assuming that the actual speed of a small ensemble of vehicles is obtained from the equilibrium speed-density relationship after retardation time, T , at an anticipated location, $x + \Delta x$:

$$v(x, t + T) = U_e(k(x + \Delta x, t)). \quad (2.4)$$

The treatment of this recursive equation is shown in detail in Muller-Krumbhaar (1987). Expanding the above equation in Taylor series with respect to T and Δx (both the quantities are kept small) yields the following equation:

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = \frac{(U - v)}{T} - \frac{C_0^2}{k} \cdot \frac{\partial k}{\partial x}, \quad (2.5)$$

where:

U is the equilibrium speed obtained from equilibrium speed-density relationship,

T is the relaxation/retardation time, and

C_0^2 is some constant interpreted as anticipation coefficient.

This relationship is discussed in detail in Section 2.3.

The difficulties arising from higher-order models, such as inability to properly describe bottlenecks and stop-start behavior, are reviewed by Hauer and Hurdle (1979). These difficulties are solved by adding a viscosity term to the momentum equation. The momentum equation, with the viscosity term, is as follows:

$$\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} = \frac{(U - v)}{T} - \frac{C_0^2}{k} \cdot \frac{\partial k}{\partial x} + \frac{v_0}{k} \cdot \frac{\partial^2 v}{\partial x^2}, \quad (2.6)$$

where v_0 is the viscosity coefficient.

The use of viscosity models is not yet widespread since its significance has not been completely understood. From an analytical point of view, the numerical viscosity term is introduced to describe bottleneck behavior and stop-start waves. In some cases, these problems can be related to the numerical methods used for the implementation of higher-order models (Michalopoulos et al. 1991).

The works mentioned above deal with a freeway as a single pipe. Some work has been done in the context of multi-lane flows (Dressler, 1949; Gazis et al 1962). This has been done by applying a simple continuum model for describing flow along two or more one-directional lanes. Changing between lanes is represented by the generation or loss of vehicles in the lane under consideration. The generation/loss term was obtained from the assumption that the exchange of vehicles between two neighboring lanes is proportional to the difference in their densities (Gazis et al 1962).

2.3 EXTENSIONS OF THE CONTINUUM MODEL

This section presents the complete mathematical formulation of model extensions applied to the higher-order continuum model. For the sake of completeness of the presentation, all three fundamental equations that constitute higher-order continuum model are briefly discussed. More emphasis is given to the model extensions.

By the analogy to fluid flows, traffic is often described in terms of flow, density, and speed. By the same analogy, we can state that a traffic flow must be conserved. This is expressed by the conservation equation. In addition to this equation, momentum equation is considered in order to account for acceleration/deceleration and inertia characteristics (Whitham, 1974; Payne, 1979). The higher-order model referred to in this research includes the basic relationship between flow characteristics, conservation equation, and the momentum equation. Speed, flow, and density of traffic at any point and instant of time can be derived by solving the system of these three equations. The three traffic characteristics define the state of the traffic system, from which measures of effectiveness such as delays and total travel time can be calculated to evaluate the performance of the system.

Sections 2.3.1 through 2.3.3 present the three fundamental equations that constitute the model. Further, Section 2.3.4 presents the extensions of the model that incorporate lane change behavior of drivers and Section 2.3.5 presents the calculation of desirable speed

for the mixed stream of drivers changing lanes and drivers continuing on the same lane. The lane change model incorporates drivers' response to the traffic control, perceived traffic, and geometric conditions. The proposed modifications elaborate the lane change process as a combination of demand for lane change and opportunity for this maneuver. The logit and queuing models are used.

2.3.1 Basic Relationship

The fundamental relationship obtained from the definitions of q , k , and v for lane i is as follows:

$$q_i = k_i \cdot v_i \quad (2.7)$$

where q_i is flow, k_i is density, and v_i is speed. The quantities q , k , and v are functions of time and space.

2.3.2 Basic Continuity Equation

The continuity equation expresses the law of conservation of traffic stream. Lanes are considered as separate pipes with an interaction between the adjacent pipes. If the traffic sinks and sources exist within the section of the freeway lane i , the continuity equation takes the following form:

$$\frac{\partial q_i}{\partial x} + \frac{\partial k_i}{\partial t} = S_{i+1,i} + S_{i-1,i} - S_{i,i+1} - S_{i,i-1} + g_i \quad (2.8)$$

where g_i is the generation/dissipation rate of vehicles at traffic sources or sinks (entrances or exits). The term $S_{i+1,i}$ represents the number of drivers who change lanes from lane $i+1$ to lane i per unit length per unit time. All the terms are functions of space and time.

The above equation implies that the total number of vehicles entering a freeway segment is equal to the total number of vehicles leaving the segment plus the storage. The left-

hand side of the equation represents the rate of change of the number of vehicles caused by the variations in flow rate and density. The right-hand side of the equation involves the lane-changing rate from adjacent lanes, entrances, and exits, which are physical sources of the change in the number of vehicles. The term $S_{i,i+1}$ is further elaborated in Section 2.3.4.

2.3.3 Momentum Equation

The momentum equation accounts for the acceleration/deceleration and inertia characteristics of the traffic mass. This equation is applicable to non-equilibrium flows. The equation for lane i is as follows:

$$\frac{\partial v_i}{\partial t} = -v_i \frac{\partial v_i}{\partial x} + \frac{(U_i - v_i)}{T} - \frac{C_0^2}{k_i} \cdot \frac{\partial k_i}{\partial x} + \frac{v_0}{k_i} \cdot \frac{\partial^2 v_i}{\partial x^2}, \quad (2.9)$$

where U_i is the equilibrium speed and T is the relaxation time. Typically, the value of U_i is calculated using an empirical speed-density relationship developed from field data.

The left-hand side of the above equation gives the acceleration/deceleration rate of the flow observed at location x . The acceleration/deceleration rate is a sum of several components:

1. velocity variations over space (first term),
2. travelers' responses to the current speed v_i different from the desirable speed U_i , (second term),
3. drivers' reactions to their own anticipation of the near-future speeds based on the observed traffic density in front of them (third term),
4. viscosity term (fourth term).

The relaxation time T , summarizes the characteristic times for acceleration/deceleration time to adjust the speed from v_i to U_i . The desirable speed U and its extensions are

discussed in the Section 3.5. The anticipation coefficient C_0^2 has the meaning of the square of the congestion velocity. This interpretation is discussed in detail by Backer (1983). The last term has no clear traffic related meaning and is postulated by some authors.

The following section presents the estimation of lane change rate mathematically.

2.3.4 Lane Change Rate

Lane change behavior of drivers is modeled for a case where a two-lane freeway without entrance and exit ramps within the vicinity of the lane closure is considered. Drivers change lanes under three different conditions.

1. Drivers comply with the DO NOT PASS signs.
2. Drivers perceive the conditions in the other lane better than in the current lane.
3. Drivers are compelled to switch since they are close to the lane drop.

Let $P_{i,i+1}$ be the likelihood that a given driver decides to change lanes during a short interval (let's say, one second long) ending at instant t . We assume that the driver makes this decision based on what he observed recently, and that this decision is independent of the decisions made during the same short interval by other drivers. The number of drivers on a unit length of single lane who decide to change lanes during the short interval can be approximated with the product $k_i \cdot P_{i,i+1}$, where k_i is traffic density in this lane at time t , and $P_{i,i+1}$ the likelihood already defined.

The lane change maneuver cannot be performed immediately since a driver has to find a sufficiently long gap in the adjacent lane and then perform the maneuver. If the time required for the lane change maneuver is $T_{i,i+1}$, then the lane change rate is given by:

$$S_{i,i+1} = \frac{k_i \cdot P_{i,i+1}}{T_{i,i+1}} \quad (2.10)$$

The next two sections describe how to calculate $P_{i,i+1}$, and $T_{i,i+1}$ based on geometry and traffic characteristics.

2.3.4.1 Proportion of Drivers Changing Lanes

Likelihood $P_{i,i+1}$ reflects drivers' willingness to change lanes in response to the perceived geometry conditions, traffic conditions, traffic control, and other factors. This section proposes a theoretical framework for a two-lane case with one lane closed. The logit model proposed in this section is suitable to incorporate all the factors into a driver's decision. The formula derived incorporates three types of factors, i.e., traffic control, geometry conditions, and traffic conditions.

Let us consider a two-lane freeway with one lane closed. Drivers approaching the lane drop select the lane they want to use. If the selected lane is the same as currently used, no lane change demand occurs. The decision (selection) process has three phases:

- (1) In the first phase, a driver perceives the DO NOT PASS message and makes the decision whether or not to comply with the message. This likelihood can be interpreted as a control compliance rate. We are assuming that drivers are consistent in their decisions to comply with the DO NOT PASS directive. It means that if a driver on the dropped lane decides to change to the continuous lane, he does not later change his decision. Similarly, if a driver on the continuous lane decides to ignore the DO NOT PASS message, he does not change his decision.
- (2) In the second phase, a driver makes the decision whether the closed lane ahead is still an alternative, given the distance to the lane's end. Drivers on the dropped lane decide to switch and drivers on the continuous lane decide to stay in the current lane if they decide to not use the discontinued lane. In the two-lane case, it is equivalent to selection of the continuous lane.

- (3) In the third phase, drivers compare traffic conditions on both the lanes (assuming that the closed lane has not been excluded yet). The lane with better conditions is selected.

The above three-stage decision process is depicted as a decision tree in Figure 2.3.1. The following notation applies:

- p^c = likelihood that a closed lane is eliminated from future use under the influence of control.
- p^g = likelihood that a closed lane is eliminated from future use based on the distance to the lane drop (geometry effect).
- $p_{i,c}^t$ = likelihood of choosing the continuous lane based on the traffic conditions (current lane as i).
- $p_{i,d}^t$ = likelihood of choosing the discontinuous lane based on the traffic conditions (current lane as i).

The likelihood of selecting a particular lane calculated based on the decision tree is as follows:

$$P_{i,d} = (1 - p^c) \cdot (1 - p^g) \cdot p_{i,d}^t \quad (2.11)$$

$$P_{i,c} = p^c + (1 - p^c) \cdot \{ p^g + (1 - p^g) \cdot p_{i,c}^t \} \quad (2.12)$$

where :

$P_{i,c}$ is the likelihood of choosing the continuous lane when the current lane is i ,

$P_{i,d}$ is the likelihood of choosing the dropped lane when the current lane is i .

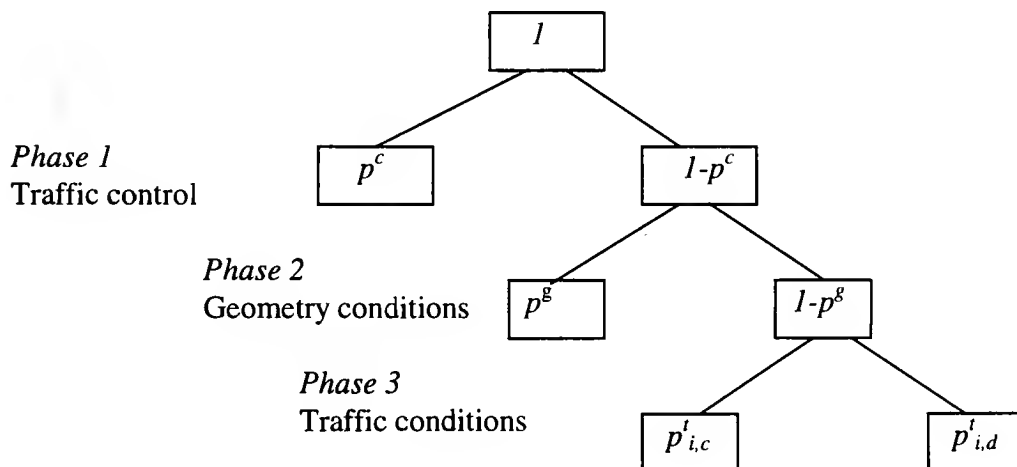


Figure 2.3.1 Decision tree for lanes change.

The likelihood p^c as postulated will give the control compliance rate. Let us assume that the compliance rate when the drivers pass the first control board is given by p^c_o . It is assumed that the pressure on the driver to comply with the control board increases as the number of control boards he passes increases. This relationship is postulated to be linear. Therefore, p^c is given by :

$$p^c = p^c_o + \alpha \cdot (n - 1) \quad (2.13)$$

where α is the rate of increase in the compliance rate, which is a measure of pressure on the drivers, and n is the number of active boards the driver passes. The control compliance rate, p^c_o and the parameter α are calibrated from the real data collected in the work zone with control boards in place.

The likelihood p^g depends on the distance between the driver and the point of the lane drop. The likelihood should take a value of one when the driver is at the point of the lane drop and near zero when the driver is far away from the lane drop. The exponential function meets these requirements and is postulated:

$$p^g = e^{-a l} \quad (2.14)$$

where l is the distance from the driver to the point of lane drop, and a is the geometry parameter to be calibrated. The value of a affects the rate of change in p^g with the distance l .

The likelihood p^l_{ij} applies to the third stage in the selection process after the driver has already decided whether to disregard the dropped lane in the further selection. The driver chooses a lane for further travel based on the traffic conditions in both lanes. Since density is a good measure of traffic conditions, it is chosen as the factor that the driver considers in the decision process. A logit model is used to estimate this likelihood:

$$p^l_{ij} = \frac{e^{-\beta(k_j + \delta)}}{e^{-\beta(k_i)} + e^{-\beta(k_j + \delta)}} \quad (2.15)$$

The parameter δ can be interpreted as the risk associated with the lane change, expressed in the equivalent density. Since this requires an attenuation effect on likelihood, δ is added to the density in the other lane and not to the density in the lane currently used by the driver. The parameter β accounts for the driver's perception of the traffic conditions. The parameters β and δ are calibrated from the real data.

2.3.4.2 Time for Lane Change Maneuver

Time, $T_{i,i+1}$, is measured between the time when the decision to change lanes is made and the time when the maneuver is completed. $T_{i,i+1}$ has two components: T_d and T_f . The first component, T_d , is the time drivers need to find a sufficiently long gap in the adjacent stream. By the analogy to the queuing theory, this term can be considered as a lane change delay. This time is strongly influenced by the opportunity for lane change. The second component, T_f , is a fixed time required to complete the maneuver after the sufficient gap is found.

The opportunity for lane change depends on the number of gaps in the adjacent lane that are sufficiently long to be accepted. The lane change can be initiated only if two conditions are met.

1. The distance between vehicles in the adjacent lane is longer than the critical space gap g_c (acceptable minimum gap).
2. The corresponding time gap is longer than the critical time gap, τ_c .

For the purpose of this model, we are assuming that the time and space gaps are distributed according to the negative exponential distribution. The analogy between pedestrians crossing the vehicle stream at unsignalized intersections and the vehicles changing lane allows us to use the modified pedestrian delay model (TRB, Special report 1975) to estimate T_d . In both cases, queuing does not occur and waiting for an acceptable gap is the only source of delay. The following modifications are, however, necessary.

(1) A driver observes the gaps in the adjacent lane moving at some speed. Thus, the average time gap has to be calculated using the relative speed Δv .

$$\text{Average relative time gap} = \frac{1}{k_{i+1} \times \Delta v_{i,i+1}}. \quad (2.16)$$

(2) The critical space gap can be converted into the equivalent time gap as follows:

$$\tau_e = g_c / \Delta v_{i,i+1}. \quad (2.17)$$

(3) The critical time gap used in the model is the larger value of τ_c and τ_e .

Finally,

$$T_d = \frac{1}{\Delta q_{i+1}} \times \frac{1}{e^{-\Delta q_{i+1} \times \tau}} - \frac{1}{\Delta q_{i+1}} - \tau \quad (2.18)$$

where $\tau = \max(\tau_c, \tau_e)$, and

$$\Delta q_{i+1} = k_{i+1} \times \Delta v_{i,i+1} \quad (2.19)$$

Relative volume Δq_{i+1} in the adjacent lane is observed by the driver moving at the relative speed $\Delta v_{i,i+1}$.

2.3.5 Desirable Speed

A desirable speed is the speed drivers consider safe and convenient under given traffic and geometry conditions. The presented model considers three factors of the desirable speed:

1. traffic density (traffic conditions)
2. lane change decision (seeking the acceptable gap)
3. distance to the lane drop (geometry condition)

Traffic density implies an average spacing between vehicles. This spacing is associated with a speed that is considered safe. Usually an equilibrium speed u_k given by the speed-density relationship is used. The speed-density relationship depends on factors such as

the width of the lane or the type of drivers. The equilibrium speed-density relationship can be calibrated from real data.

After the lane change decision is made, the driver starts seeking a suitable gap to merge. The speed at which the driver is traveling substantially affects the expected maneuver delay. It can be shown that the optimal relative speed minimizing the lane change time is equal to g_c/τ_c , where g_c is the critical space gap and τ_c is the critical time gap. This means that drivers who want to change lanes tend to travel at the optimal speed higher than in the continuous lane. This behavior was observed in field data (see Section 2.5.5). The desirable speed for the drivers changing lanes, u_c , is computed as the sum of the speed in the adjacent lane and the optimal relative speed Δv .

It could be expected that drivers approaching the end of the dropped lane slow down to be able to stop the vehicle in case the merging maneuver is not possible, but field data indicate that if the continuous lane is filled with sparse traffic, drivers in the dropped lane maintain high speed. On the other hand, if the speed of traffic on the continuous lane is low, then the speed on the dropped lane near the lane drop is only slightly higher (by approximately 15 km/h). This means that the geometry impact on speed need not be explicitly considered for desirable speed calculations. This factor is indirectly incorporated through the other factors.

Finally, to calculate the desirable speed for the freeway segment, the drivers are divided into two groups: drivers who have decided to change lanes and drivers who have decided to stay in the same lane. The value of the desirable speed is calculated as the expected value. Specifically, if P_c is the fraction of drivers who have decided to change lanes, then the desirable speed U is given by an equation:

$$U = P_c \cdot u_c + (1 - P_c) \cdot u_k \quad (2.20)$$

2.4 MODEL DISCRETIZATION

As mentioned in Section 2.2.2.1, solving a higher-order model for realistic boundary conditions is difficult. This chapter describes briefly how the second-order model presented in the previous chapter is solved numerically by discretizing both time and space. The discretized model presented in this chapter is for a typical rural case of two-lane freeway with no entrance and exit ramps along the section approaching a work zone.

2.4.1 Background

The model consists of three equations, two of which are partial differential equations involving time and space. It is very difficult to obtain a closed-form solution to compute flow, speed, and density at any point of time and space for realistic boundary conditions. The practical way of solving these equations is the numerical method. However, simple forward discretization schemes are not suitable, and in many cases lead to wrong results due to the numerical instabilities. The numerical methods must contain (1) an implicit integration procedure with centered differences, and (2) correct use of the boundary and initial conditions.

To implement the numerical methods, each lane of the freeway segment is divided into a number of sections as shown in Figure 2.4.1. Density and speed are attributed to the entire section, while flow is attributed to the end of the section. An outflow from one section is the inflow to the following section (Figure 2.4.2).

The correct choice of the spatial and temporal step size for discretization of the space and time coordinates is crucial for the numerical solution. The spatial step, Δx and the temporal step, Δt are connected via the characteristic speed V . The relationship is as follows:

$$V \cdot \Delta t \approx \Delta x \quad (2.21)$$

Experience with numerical solutions shows that sections and time intervals have to be sufficiently small (Papageorgiou et al. 1990). The larger the step size, the less accurate are the results but the less time it takes to complete the simulation. The smaller the step size, the more accurate are the results but the more time it takes for simulation. So, there is a compromise in choosing the step size based on the desirable accuracy of the results and time of simulation. For the current implementation $\Delta x = 50$ m and $\Delta t = 1$ sec are chosen.

The next three sections, 2.4.2 through 2.4.4, present the transformation of the continuum model into a system of three equations with unknown variables $k_{i,j}^{t+1}$, $V_{i,j}^{t+1}$, and $q_{i,j}^{t+1}$ where:

$k_{i,j}^{t+1}$ = density of traffic in lane i , segment j and time $t+1$,

$V_{i,j}^{t+1}$ = Speed of traffic in lane i , segment j at time $t+1$, and

$q_{i,j}^{t+1}$ = outflow rate from segment j in lane i at time $t+1$.

2.4.2 Basic Relationship

The basic relationship in discretized form for lane i , and segment j , and for time interval $t+1$ is as follows:

$$q_{i,j}^{t+1} = k_{i,j}^{t+1} \cdot V_{i,j}^{t+1} \quad (2.22)$$

Since density and speed are attributed to the entire section, the basic relationship in the discretized form gives the outflow from a section. Outflow from the previous section $j-1$ will be inflow to the next section j in the same lane.

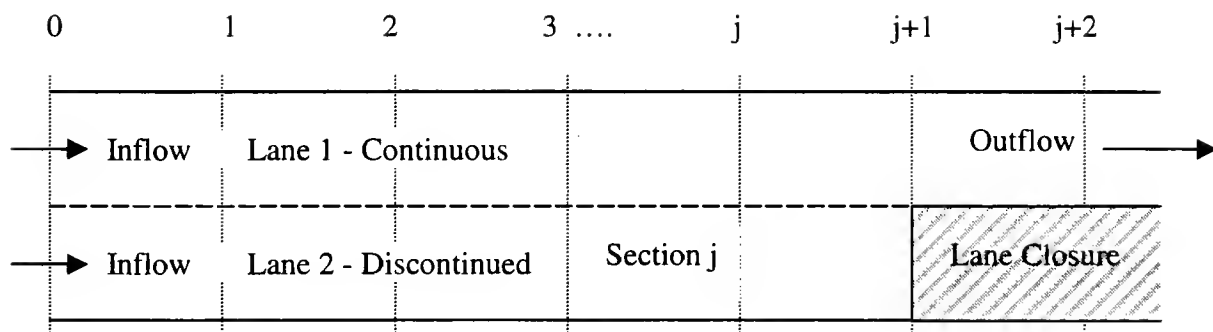


Figure 2.4.1 Sample Layout of a Two-lane Freeway with One Lane Dropped

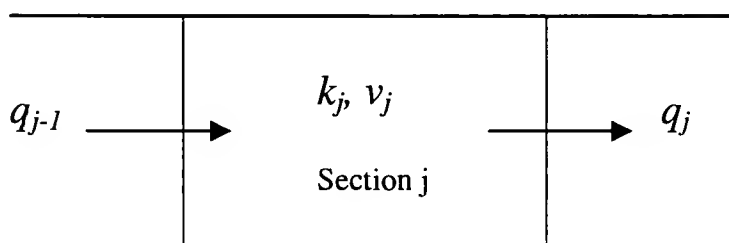


Figure 2.4.2 A Typical Section of a Freeway Lane

2.4.3 Basic Continuity Equation

The following is the discretized form of the basic continuity equation:

$$\frac{k_{ij}^{t+1} - k_{i,j}^t}{\Delta t} + \frac{q_{i,j}^t - q_{i,j-1}^t}{\Delta x} = (S_{i,k}^{t,j} - S_{k,i}^{t,j}) \quad (2.23)$$

and

$$k_{i,j}^{t+1} = k_{i,j}^t - \frac{\Delta t}{\Delta x} (q_{i,j}^t - q_{i,j-1}^t) + \Delta t (S_{i,k}^{t,j} - S_{k,i}^{t,j}) \quad (2.24)$$

The updated density, $k_{i,j}^{t+1}$, is computed assuming that flow rates, and lane change rates remain constant over the short interval Δt between the instants t and $t+1$. The term $\frac{\Delta t}{\Delta x} \cdot (q_{i,j}^t - q_{i,j-1}^t)$ gives the average change in density in the time interval Δt between instants t and $t+1$ due to the unbalance inflow $q_{i,j-1}^t$ and outflow $q_{i,j}^t$. The term $\Delta t \cdot (s_{i,k}^{t,j} - s_{k,i}^{t,j})$ is the change in density due to lane change in the time interval Δt between instants t and $t+1$.

2.4.4 Momentum Equation

The following is the discretized form of the traffic propagation equation.

$$\frac{V_{i,j}^{t+1} - V_{i,j}^t}{\Delta t} + \frac{V_{i,j+1}^t - V_{i,j-1}^t}{2 \Delta x} \cdot V_{i,j}^t = \frac{U_{i,j}^t - V_{i,j}^t}{T} - C_o^2 \cdot \frac{k_{i,j+1}^t - k_{i,j}^t}{\Delta x k_{i,j}^t} \quad (2.25)$$

and

$$V_{i,j}^{t+1} = V_{i,j}^t + \left\{ \frac{U_{i,j}^t - V_{i,j}^t}{T} - \frac{V_{i,j+1}^t - V_{i,j-1}^t}{2 \cdot \Delta x} \cdot V_{i,j}^t - C_o^2 \times \frac{k_{i,j+1}^t - k_{i,j}^t}{\Delta x \cdot k_{i,j}^t} \right\} \cdot \Delta t \quad (2.26)$$

In the above equation, the term $\frac{U_{i,j}^t - V_{i,j}^t}{T}$ is the acceleration applied by drivers at time t to reach the desirable speed $U_{i,j}^t$. The term $\frac{V_{i,j+1}^t - V_{i,j-1}^t}{2\Delta x} \cdot V_{i,j}^t$ is the acceleration caused by spatial changes in speed, and the term $C_o^2 \times \frac{k_{i,j+1}^t - k_{i,j}^t}{\Delta x \cdot k_{i,j}^t}$ is the extra acceleration/deceleration applied by the drivers in anticipation of the near future traffic conditions based on the density observed ahead.

2.4.5 Boundary Conditions

As can be seen from Equation 2.24, the density in section j at time $t+1$, $k_{i,j}^{t+1}$, depends on the flow from the previous section $j-1$ at time t denoted by $q_{i,j-1}^t$. Similarly, the speed in section j at time $t+1$, $V_{i,j}^{t+1}$, depends on the speed in the previous section $j-1$ at time t , $V_{i,j-1}^t$. These values for the first segment, $V_{i,0}^t$ and $q_{i,0}^t$ for every lane i , must be given as the boundary values at every instant t .

One other type of boundary conditions is needed to simulate the lane blockage. If a lane is discontinued, the flow rate beyond that point is zero, as well as the lane changes to the discontinued lane. Therefore, if lane i is dropped after m sections, counting from the entry section (as shown in Figure 2.4.1), then $q_{i,m}^t$ and lane change rates are set to zero for all times t , beyond the section m .

2.4.6 Initial Conditions

From Equations 2.22, 2.24, and 2.26, it can be seen that the values of density, flow, and speed at each instant depend on the values of density, flow, and speed at the previous instants. These values must be given for the first instant. The values should be reasonable because of their strong impact on the following traffic states.

In order to minimize the bias in a few initial iterations, simulation must begin with some reasonable initial conditions. The initial conditions are obtained by applying a warm-up time (about 10 minutes). The warm-up simulation starts with free-flow conditions. Density, flow, and speed are computed during the warm-up time with all the boundary conditions set as discussed in the previous section. The inflow and speed in each interval of the warm-up time are equal to the input values for the first interval of the actual simulation period.

2.4.7 Solution Approach

After the initial conditions are obtained, the values of density, flow and speed at every point are updated for every instant as follows.

First, the lane changes rate, and desirable speed for every section is calculated. Then, density is updated for the next instant $t+1$ based on the values at the instant t , using the modified basic continuity equation 2.24. Once density is determined, the speed at time step $t+1$ is obtained based on the modified momentum equation 2.26. Lastly, flow is calculated for the time step $t+1$ based on Equation 2.22.

This procedure is repeated for all sections and for each instant. The methodology of calculating flow speed and density at each time step is summarized in flowchart given in Figure 2.4.3. The above described discretized model has been coded in C++ for running the simulation.

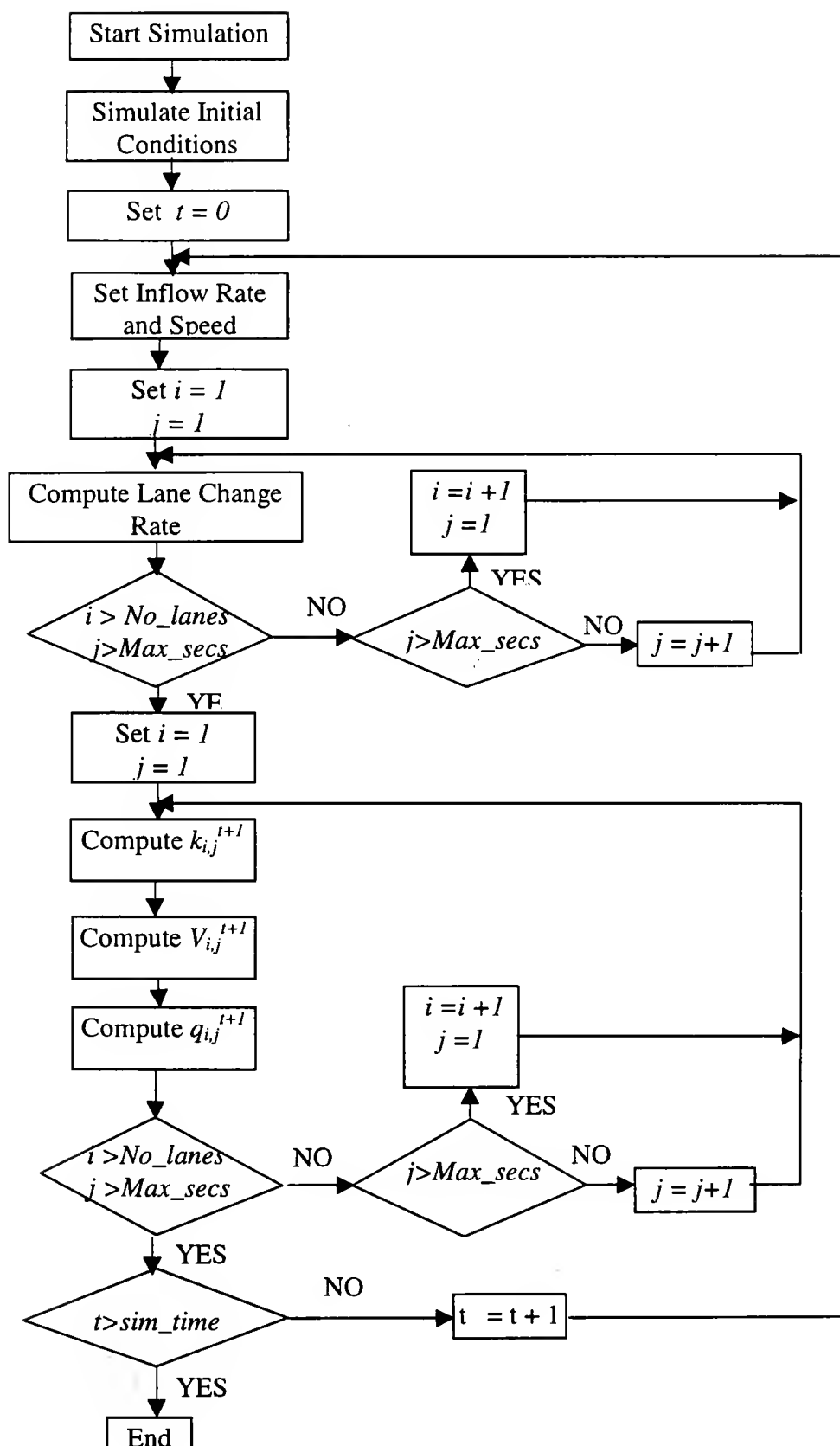


Figure 2.4.3 Flowchart for Calculating Flows, Speeds, and Densities

2.5 DATA COLLECTION

This section briefly describes the methods adopted for data collection and data processing and presents the results from the field data.

2.5.1 Site Selection

A selection of suitable work zones for the data collection is crucial. Initially, the information on maintenance and reconstruction activities planned or taking place in the state of Indiana was requested from the Indiana Department of Transportation. Then, project location, type of work zone, number of lanes closed, length of the work zone, and chance for congestion was considered. The work zone appeared feasible if the following conditions were met.

- The work zone has to be reasonably long (at least about one-half mile) for drivers to adjust to a desirable speed in the work zone. This helps in measuring the capacity of the work zone.
- The freeway should have heavy traffic so that the flow reaches the capacity of the work zone and queues are formed. This is important because the driver's compliance to control becomes more critical when the flow in the work zone is at its capacity.
- There should not be any on/off ramps for about 3 to 5 km in the approach zone and also no on/off ramps within the work zone.

After examining the work zone projects planned or conducted in Indiana in 1996, two freeway work zones on I-69 and on I-74 near Indianapolis were selected. Traffic congestion was expected on the I-69 southbound roadway during the morning peak and during the evening peak on the I-69 northbound roadway. Traffic congestion was expected during the morning peak on the I-74 westbound roadway. Unfortunately, the I-74 eastbound roadway had an entrance ramp in the work zone and an exit ramp close to the entry point of the work zone. So the I-74 eastbound roadway location was not

selected. Finally, the I-69 southbound roadway, the I-69 northbound roadway, and the I-74 westbound roadway were the three sites selected for the data collection.

2.5.2 Techniques Applied in Data Collection

The purpose of the data collection was to calibrate and evaluate the developed model. For calibration and evaluation, it was desirable to define a period when the freeway was initially uncongested and then was congested. This would ensure a proper modeling of the transition from uncongested to congested traffic conditions on a freeway. In order to find the approximate time of peak flow, the flow variation on the freeway was observed for two or three days before the actual day of data collection. It was found that the I-69 southbound freeway and I-74 westbound freeway had a peak flow about 7:30 to 9:30 a.m. and the I-69 northbound freeway had peak flow about 4:30 to 6:00 p.m.

Once the peak flow times were determined, flow rate and speed of traffic were measured at several selected points using magnetic detectors (Nu-Metric detectors shown in Fig. 2.5.1 and Fig 2.5.2). These detectors were placed at least one hour before the peak period started to measure the traffic conditions preceding the peak flow occurrence. Distance Measuring Vehicle (DMV) was used to measure the distances between the detectors with one-foot accuracy.

The following Sections 2.5.2.1 and 2.5.2.2 describe how the detectors and DMV were used in the data collection.

2.5.2.1 Nu-Metrics Detectors

Two types of Nu-Metrics magnetic detectors were used in the data collection. They are (1) Countcard (Fig 2.5.1), and (2) Hi-Star (Fig 2.5.2). Both types of detectors give counts (number of vehicles passed in a specified time interval) and average speeds (for the pre-

specified time interval). Countcard gives aggregate data for a minimum time interval of 15 minutes. Hi-Star gives aggregate data for a minimum interval of one minute. Traffic conditions in small intervals on the approach segment to the work zones were crucial to grasp short-term traffic fluctuations. Hi-Stars were used to collect aggregate data in five-minute intervals, while Countcards were used to measure the outflow rate (capacity of the work zone) and speed for 15 minute intervals.

To measure the traffic data, the detectors had to be programmed to set the start time and the stop time of the data collection, and the aggregation interval. Once all these settings were done, the detectors were placed on the freeway at predetermined locations as shown in Figure 2.5.3. A detector at location 4 (Figure 2.5.3) was placed inside the work zone 800 meters beyond the lane merge point. Detectors at location 3 were placed 50 meters to the lane merge point, one on each lane. Other detectors were evenly spaced at the frequency of 1,200 meters.

2.5.2.2 Distance Measuring Vehicle

It was necessary to have accurate distances from the lane merge point to all the detectors. A distance-measuring vehicle was used to collect the data at one-meter accuracy. The odometer was set at zero at the first detector. Then, the distance to the other detectors and the lane merge point were measured.

2.5.3 Data Processing

Once the required data were collected, it was processed to obtain the format required by the simulation model, which involved several steps.

First, the data were downloaded to the computer from the detectors via an interface unit using software (LP software) provided by Nu-Metrics. The data from each detector were

stored in various files in the format that could be viewed only by the LP editor. The format included start times of intervals, counts, and average speeds.

Then, the files were exported to Excel (spreadsheet) format. After exported to Excel, the vehicle counts were grouped in fifteen speed bins (each speed bin was 8 km/h long).

Finally, the Excel files were re-formatted to obtain three columns with time interval numbers, hourly flow rates, and average speeds. The files were converted to a text format to enable reading the files by the simulation program.

2.5.4 Data Collection with Control in Place

For collecting data with DO NOT PASS boards, the boards were placed at the work zone sites one or two weeks in advance so that drivers could become familiar with the new control system. The intention was to reduce the bias in the driver compliance rate to the control. Due to technical difficulties, the boards were placed a few hours before the data collection, and this should be taken into account when interpreting the results.

Three DO NOT PASS boards were used in the data collection. These boards were placed on the side of the discontinued lane at locations predetermined by the INDOT specialists (Figure 2.5.4). One board was placed very close to the lane merge point and the other boards were placed evenly spaced. The board close to the lane merge point was kept turned on all the time. The other boards were activated manually by INDOT personnel every time the queue reached the location of the neighboring downstream board. The activation and deactivation times of these boards were noted, and the distances between the control boards and the lane merge point were measured using DMV.

The detectors were placed to collect data at exactly the same location as during the data collection without boards. The data processing was the same as well.

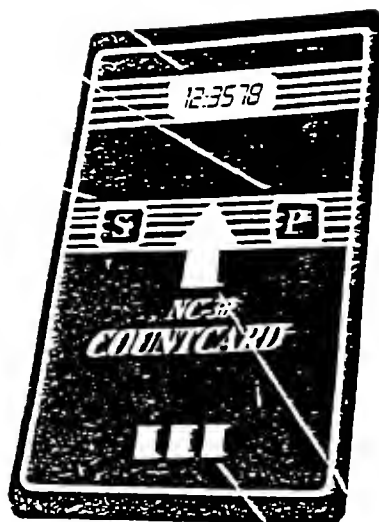


Figure 2.5.1 Countcard Detector

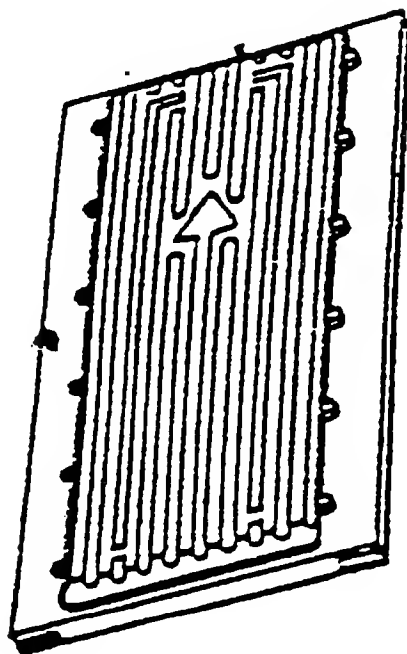


Figure 2.5.2 Hi-Star Detector

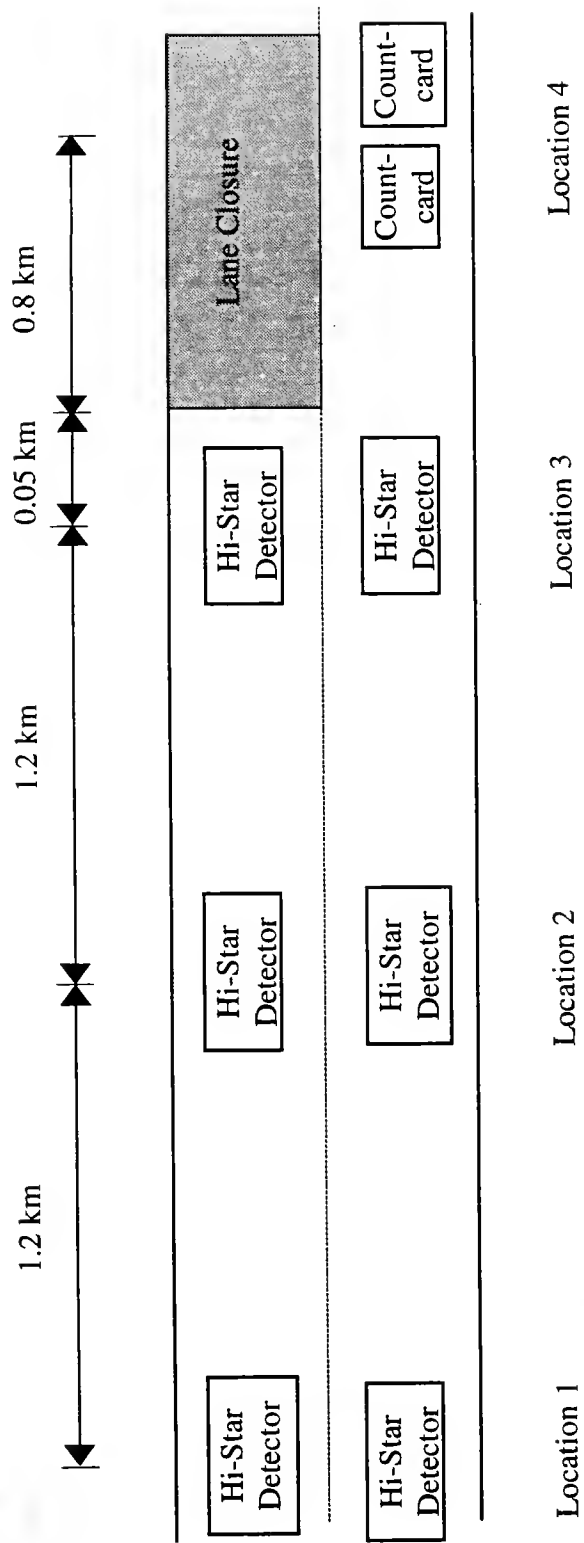


Figure 2.5.3 Layout of Detectors for Data Collection without Control

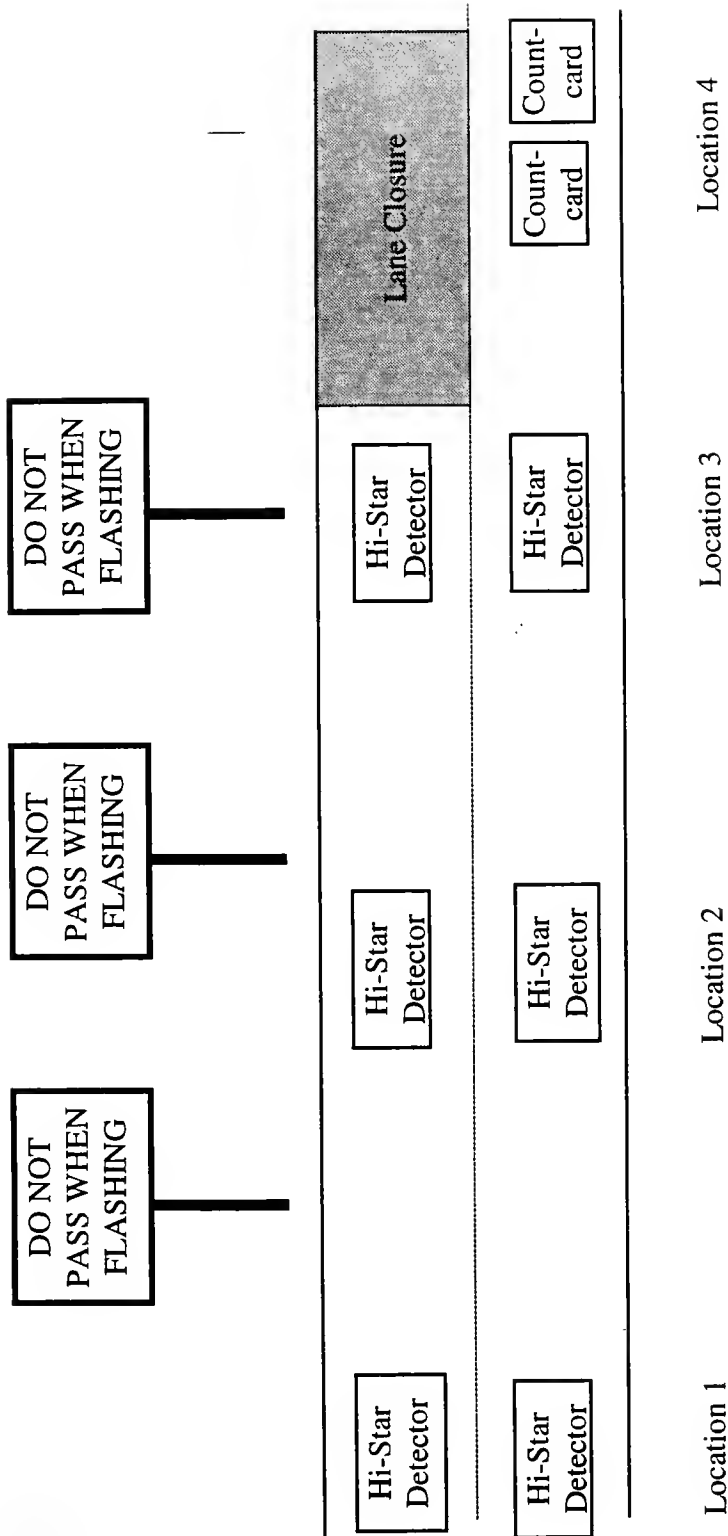


Figure 2.5.4 Layout of Detectors for Data Collection with Control

2.5.5 Discussion of the Results of Field Data Collection

After processing the data for both without and with control, graphs were plotted to make a comparative study of the variation of flow, speed, and density over time at various distances to the lane drop. Similar results were observed at all the sites. For illustration purposes, the results from I-69 southbound site are presented here.

Figures 2.5.5 and 2.5.6 indicate that the inflow rate and distribution of inflow in both lanes are similar in the both cases without and with control. This is a desirable situation if one has to compare the two cases. There is a sudden and unusual change in volume distribution between lanes in the case with control at 9:30. The results for the period after 9:30 are not considered in further discussion.

Figures 2.5.7 and 2.5.8 indicate that density in the dropped lane is decreasing as the lane drop point is approached and density on the continuous lane is increasing. This indicates that more and more drivers move from the dropped lane to the continuous lane as they get closer to the lane drop point. This can be attributed to the geometry factor. In addition, it can be seen that the traffic density in the dropped lane is lower when the boards are in place. This can be attributed to the control effect.

The higher lane change rate from the dropped lane to the continuous lane results in the flow on the continuous lane higher than in the dropped lane (Fig. 2.5.9 and 2.5.10).

From Figures 2.5.11 and 2.5.12, it can be seen that the speed in the dropped lane is consistently higher than the speed in the continuous lane. The speed difference is, in general, fixed, which supports the hypothesis that the drivers who want to change lanes from the dropped lane to the continuous lane travel at an optimal relative higher speed than drivers in the continuous lane (Section 2.3.5).

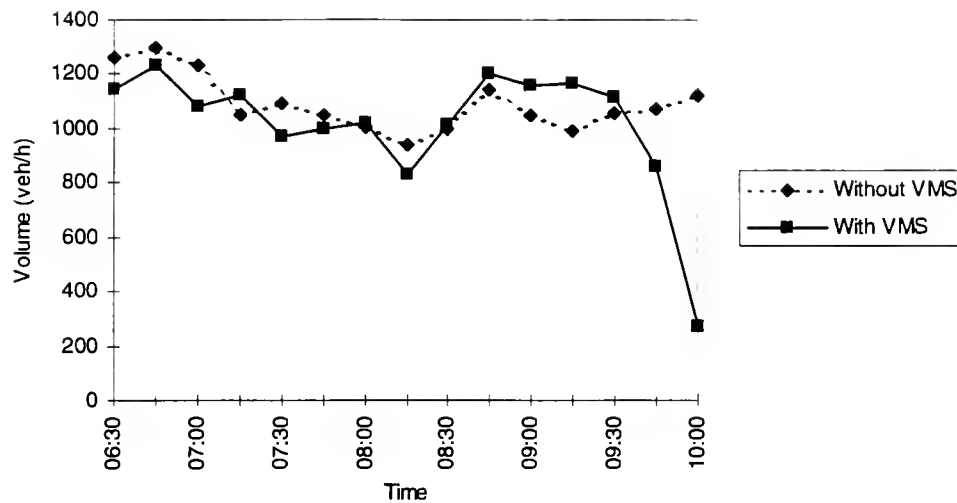


Figure 2.5.5 Variation of Total Inflow on I-69 Southbound Approach

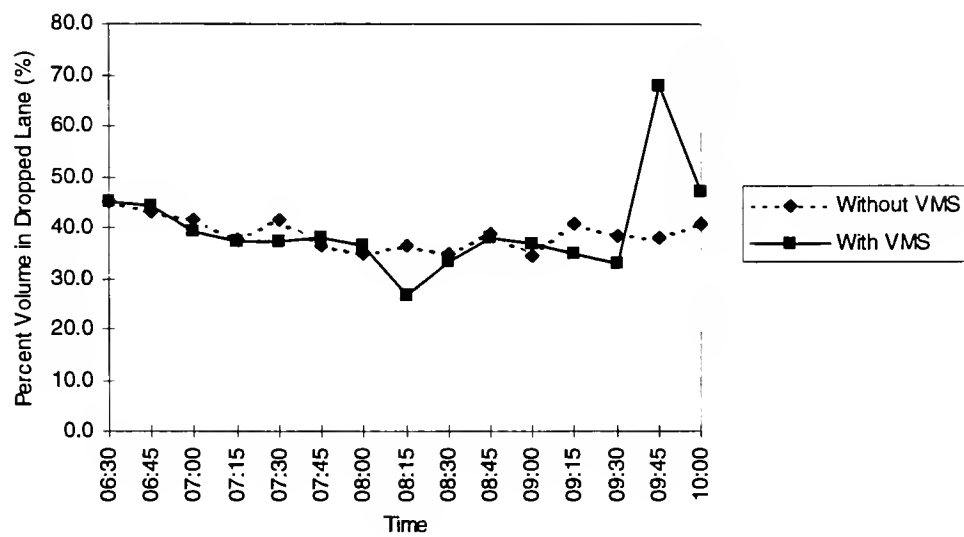


Figure 2.5.6 Variation of Traffic Distribution between Lanes on I-69 Southbound Approach

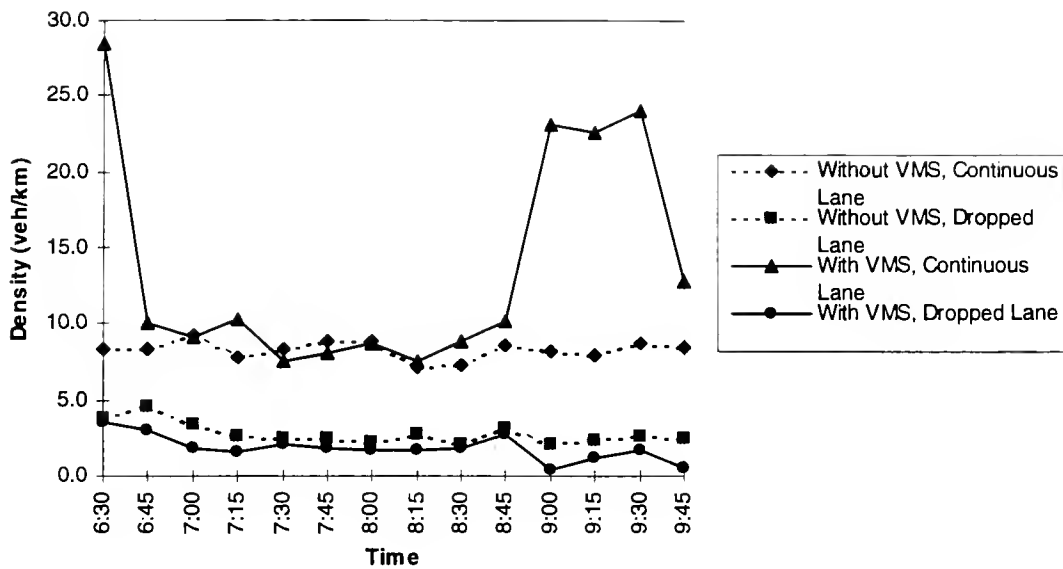


Figure 2.5.7 Density Variation 1.25 km to Lane Drop (Detector Location 2)

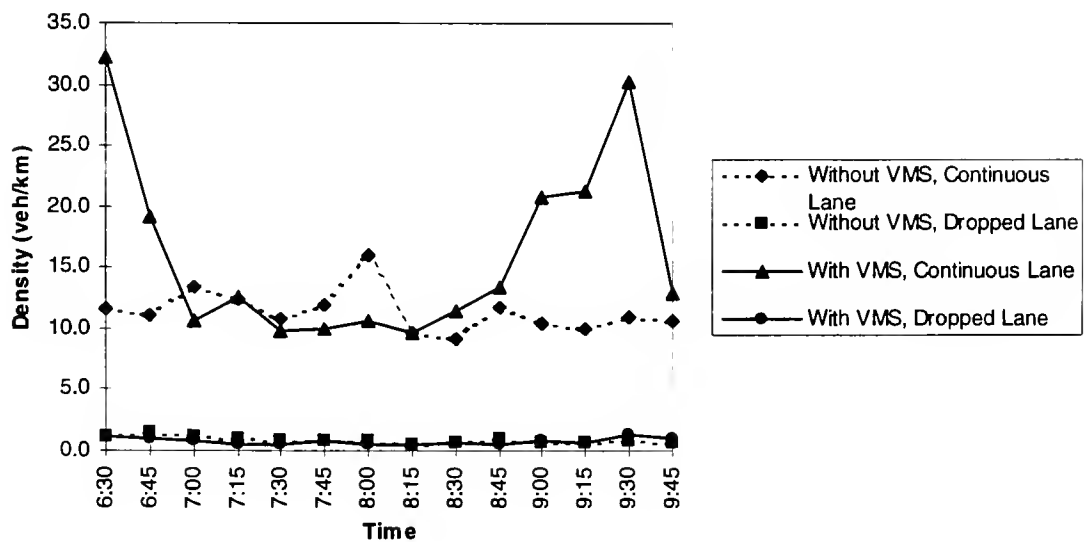


Figure 2.5.8 Density Variation with Time at 0.05 km to Lane Drop (Detector Location 3)

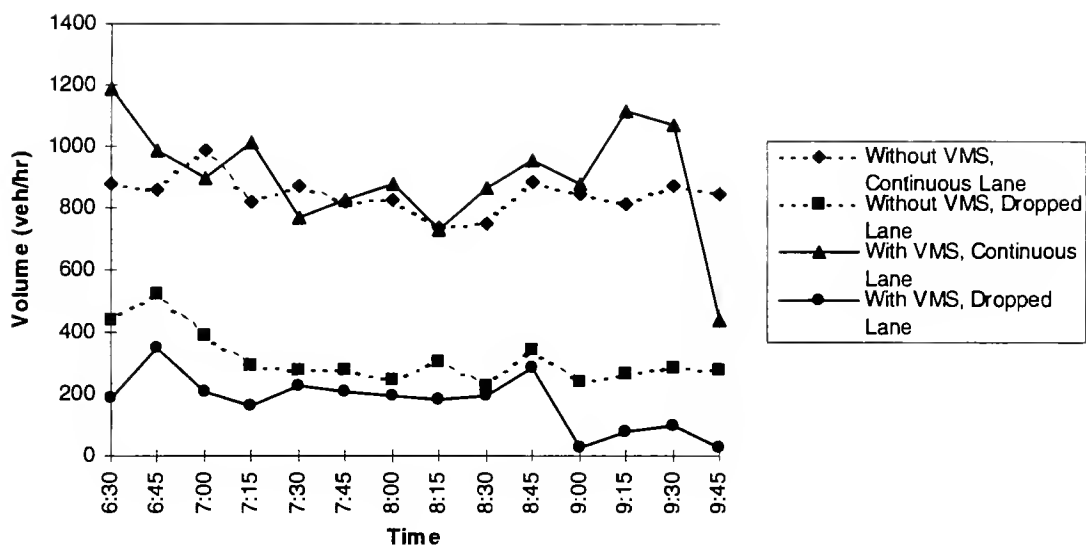


Figure 2.5.9 Flow Variation with Time at 1.25 km to Lane Drop (Detector Location 2)

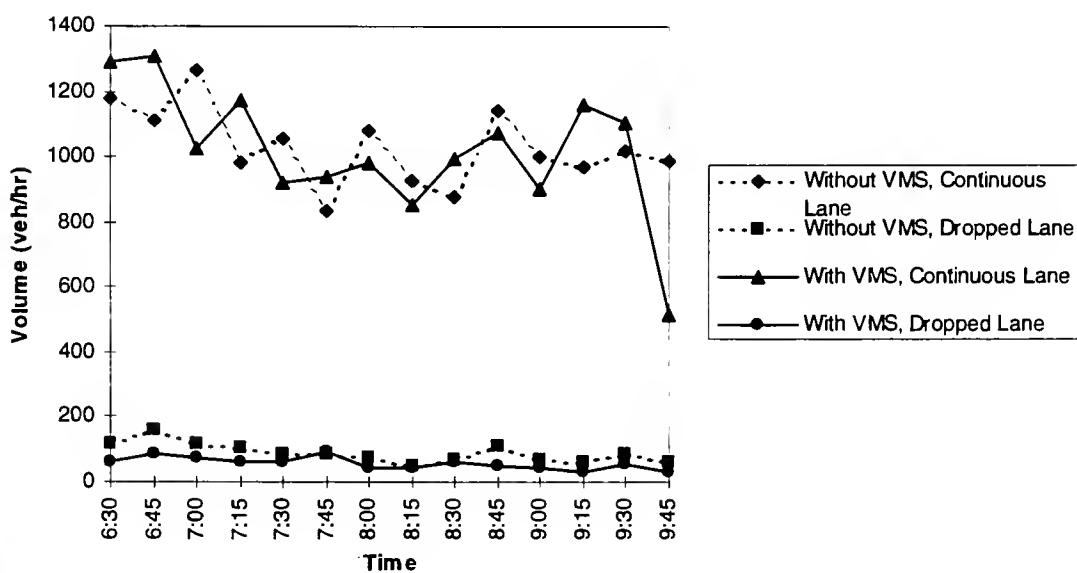


Figure 2.5.10 Flow Variation with Time at 0.05 km to Lane Drop (Detector Location 3)

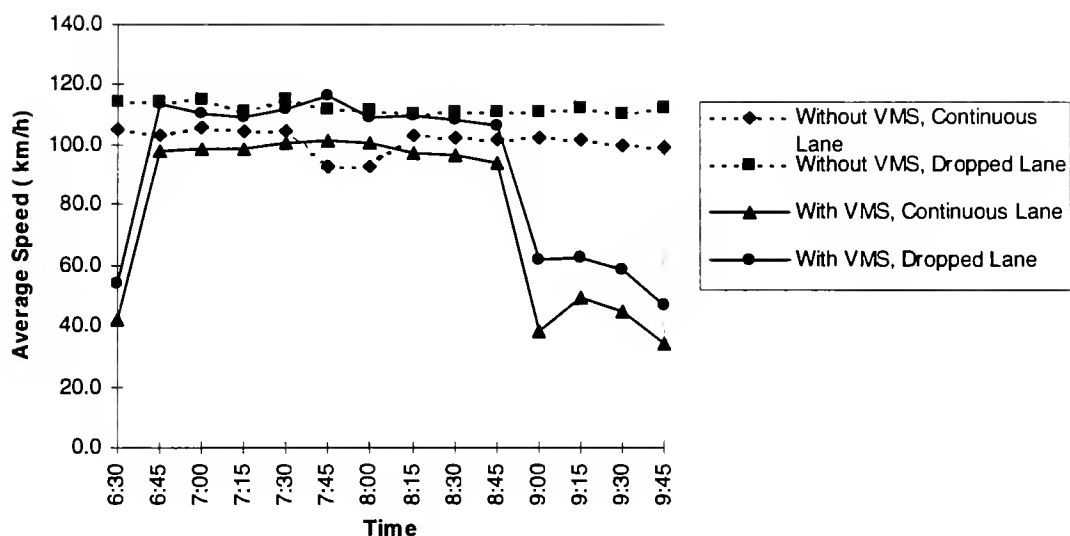


Figure 2.5.11 Speed Variation with Time at 1.25 km to Lane Drop (Detector Location 2)

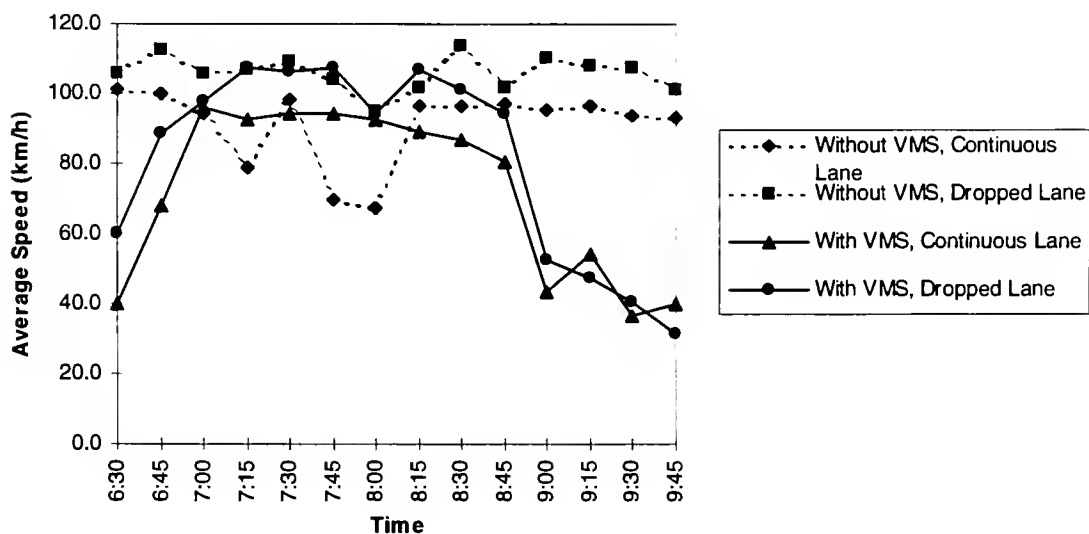


Figure 2.5.12 Speed Variation with Time at 0.05 km to Lane Drop (Detector Location 3)

2.6 CALIBRATION OF THE MODEL

Models involving parameters require calibration. In the calibration process, the parameter values are adjusted so that the discrepancy between the output from the model and the values observed in the field is minimized. This chapter presents the calibration cases and the methodology adopted for calibration. Then, the results from calibration are presented.

2.6.1 Calibration Process

Several model parameters need to be calibrated. The following gives the list of parameters that need to be calibrated in the Second-Order model proposed in section 2.3.

- Equilibrium speed-density relationship,
- Geometry factor,
- Anticipation factor,
- Viscosity factor,
- Driver's Perception factor,
- Driver's Risk factor,
- Initial Compliance Rate, and
- Incremental Compliance Rate.

The calibration of the above parameters is done in three stages. In the first stage, the equilibrium speed-density relationship alone is calibrated. This relationship does not depend on any other parameters. The actual speed is calculated using the equilibrium speed-density relationship (Equation 2.3). In the second stage, the factors independent of control are calibrated. The parameters geometry factor, anticipation factor, viscosity factor, driver's perception factor, and driver's risk factor are assumed the same in both the cases with and without control. These parameters are calibrated using the field data collected in the work zones without control. Then the control parameters, i.e., initial

compliance rate and incremental compliance rate, are calibrated using the field data collected with control boards, keeping the values of the other parameters unchanged that were calibrated in the previous two stages.

The next three sections briefly present the methodology adopted for calibration of the equilibrium speed-density relationship, and other control-independent and control dependent parameters.

2.6.2 First Stage of Calibration

As mentioned previously, in the first stage, the equilibrium speed-density relationship is calibrated, the methodology for which follows.

A scatter plot of points represented by speed - density pairs are made for each site separately as shown in Figures 2.6.1 and 2.6.2 for the I-69 freeway work zone. To capture the variability of speed-density relationship as close as possible, three regression lines are fit in three different regions of the scatter plot. The figures show that the speed almost remains at free flow speed up to a density of 6 veh/km. Beyond the density of 6 veh/km, the speed reduces rather rapidly up to a density of 31 veh/km. Beyond the 31 veh/km density, the speed reduces more slowly and becomes zero at jam density. Similar variation was observed for other sites also.

Comparison of the scatter plots for I-69 southbound and I-69 northbound freeways (Fig. 2.6.1 & 2.6.2) indicates that, though the free-flow speeds for both the sites are different, the speed-density variation beyond 31 veh/km density are almost the same. In order to generalize the speed-density relationship, the following procedure is adopted. For density less than 6 veh/km, a free-flow speed, specific for a given site, is used. For density higher than 31 veh/km, a common straight line is used for all the sites. For density between 6 and 31 veh/km, the speed-density relationship is assumed to be a straight line that

connects the two straight lines determined for the density region 0 - 6 and > 31 veh/km. The speed at a density around 31 veh/km/lane is approximately observed to be 41 km/h.

The example speed-density relationship obtained for I-69 northbound freeway is plotted in Figure 2.6.3. The free-flow speed is taken as 108 km/h. The equilibrium speed-density relationship is obtained as follows:

$$u_k = \begin{cases} 108 & \text{if } k < 6 \\ 125 - 2.697 \cdot k & \text{if } 6 < k < 31 \\ 77 - 1.173 \cdot k & \text{if } 31 < k < 66 \end{cases} \quad (2.27)$$

Using this speed-density relationship, the parameters are calibrated for the real data without and with control. The following two sections describe the next two stages of calibration.

2.6.3 Second Stage of Calibration

Once the speed-density relationship is calibrated, reasonable values for other parameters are necessary to run the simulation model. As mentioned earlier, the values of the parameters that do not account for the effect of control are set so that the discrepancy between the real data and the simulated results are minimized for data collected without using control boards. The total of squared differences in observed and simulated densities has been chosen as an error function. The error function has to be evaluated. Since it does not have a closed form, it is very difficult to evaluate the derivatives of the error function with respect to the calibrated parameters. Many of the existing optimization methods cannot be applied to this problem. The following section describes the complex method that was used for calibration.

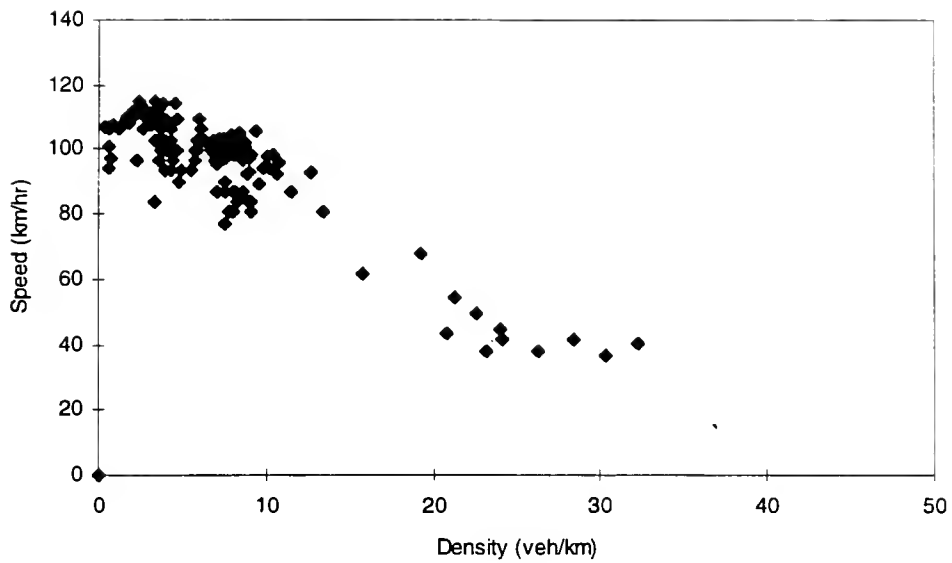


Figure 2.6.1 Scatter Plot of Density and Speed for I-69 Southbound Freeway Approach

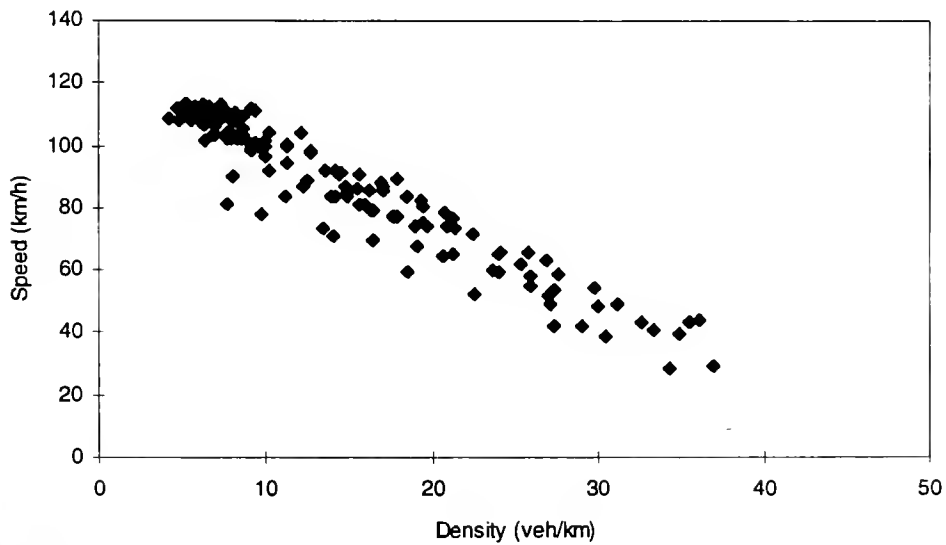


Figure 2.6.2 Scatter Plot of Density and Speed for I-69 Northbound Freeway Approach

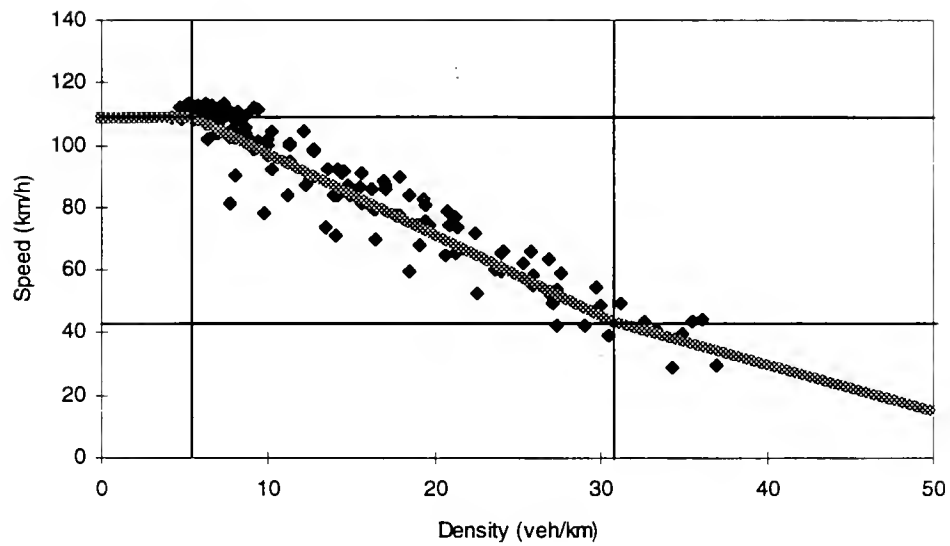


Figure 2.6.3 Proposed Relationship between Density and Speed (I-69 Northbound Freeway Data)

2.6.3.1 Method and Algorithm

The complex method, as a non-gradient search method, fits the nature of the optimization problem. The complex method does not require the derivative of the error function to find the minimum point, and hence can be applied to this problem. The complex method starts with a set of P points generated randomly and sequentially. The point is a vector of the parameter values. Given the parameter upper bound, B^u , and lower bound, B^l , the random numbers uniformly distributed in the interval (0,1) are generated, and the values of the parameters are calculated using the following equation:

$$\text{parameter value} = B^l + (B^u - B^l) \cdot \text{random number} \quad (2.28)$$

N random numbers are required to define a point with N parameters.

Given the set of P points, the error value is evaluated for each point by running the simulation. The point corresponding to the highest error value is rejected as the worst one. A new point is generated by reflecting the rejected point a certain distance through the centroid of the remaining points. Thus if X^r is the rejected point and X^c is the centroid of the remaining points, then the new point is calculated using the following relationship:

$$X^m = X^c + \nu (X^c - X^r) \quad (2.29)$$

The multiplier ν determines the distance of the reflection: $\nu = 1$ corresponds to setting the distance $\|X^m - X^c\|$ equal to $\|X^c - X^r\|$; $\nu > 1$ corresponds to an expansion; and $\nu < 1$ corresponds to a contraction.

At the new point, the error value and the bounds are checked. There are several possibilities.

1. The new point is feasible and its error value is not the highest in the current set of points. In this case, reject the worst and continue with the reflection operation.
2. The new point is feasible and its function value is the highest of the current set of P points. Rather than reflecting back again (which would cause cycling), retract the point by half the distance to the previously calculated centroid.
3. The new point is not feasible, i.e., the value of the parameter is beyond its upper or lower bound. Set the parameter value to its upper or lower bound appropriately.

The search is terminated when the pattern of points has shrunk so that the points are sufficiently close to each other and/or when the differences between the function values become small enough.

The upper and lower bounds for all the parameters are required for applying the complex method. For computational efficiency, the bounds must be as tight as possible. The complex algorithm has been coded in C++ .

The next section gives the details of the data used for calibrating the model both without and with control cases and presents the results obtained from calibration.

2.6.3.2 Calibration Results

The complex method was run a number of times with different sets of randomly generated initial feasible points with the upper and lower bounds given in the Table 2.6.1. The algorithm yielded different solutions. This means that the objective function has a many local optima (objective function is not convex). Thus, the best solution was selected. The corresponding values of geometry factor, anticipation factor, viscosity factor, driver's perception factor, and driver's risk factor are tabulated in Table 2.6.2. The geometry factor is set at value 0.001579 (distance to the lane drop is expressed in meters). This means that about 20% of the people in the dropped lane will be willing to

merge the continuous lane at a distance of 1 km to the lane drop. Risk factor of 0.0400 means that drivers who consider changing lanes based on traffic conditions consider an additional risk equivalent to 0.04 veh/m/lane (or 40 veh/km/lane). This high value indicates that the drivers consider lane changing risky and that they change lanes only when they are compelled to do so.

From Figures 2.6.6, 2.6.9, 2.6.12, 2.6.15, and 2.6.18, it can be seen that the model captures the trend in variation of density very well. In addition, the distribution of vehicles between lanes at 0.61 and 0.15 mile to the lane drop is predicted closely. On the other hand, the two sharp peaks in density observed in the real data for the continuous lane (Fig. 2.6.12) are not captured in the model. The peaks could be caused by a temporary blockage of the continuous lane by the working crew and not a queue formation due to flow reaching capacity in the work zone. This hypothesis is supported by Figure 2.6.16 where the outflow rate during the high-density period is lower when compared to other periods. Figures 2.6.4, 2.6.7, 2.6.10, 2.6.13, and 2.6.16 show that the flow variations and distribution between the two lanes are captured very well. Figure 2.6.16 indicates that the outflow rate is predicted reasonably close to the actual data. This is very critical because the work zone capacity (represented by the outflow rate) during high-density conditions controls the outflow from the segment, which in turn determines the queue lengths, travel times, etc.

Table 2.6.1 Traffic Parameters' Constraints (distance is in meters and time is in seconds)

Parameter	Lower Bound	Upper Bound
Geometry Factor	0.0000001	0.05
Anticipation Factor	1.0	15.0
Viscosity Factor	1.0	2000
Risk Factor	0.0000001	0.5
Perception Factor	50	7000

Table 2.6.2 Calibrated Traffic Parameters (distance is in meters and time is seconds)

Parameter	Calibrated Value
Geometry Factor	0.00157
Anticipation Factor	1.4342
Viscosity Factor	23.51
Risk Factor	0.040
Perception Factor	4793.13

From Figures 2.6.5, 2.6.8, 2.6.11, 2.6.14 and 2.6.17, it can be observed that the average speed variation is also very well captured at various distances to lane drop and in both lanes. In Figures 2.6.5, 2.6.11, and 2.6.14, there are two sharp reductions in average speed, which are not captured in the simulation. The reason could be the same as mentioned for the density case above.

In general, we can state that the model produces results that closely follow the actual data. This is a strong indication of the model's robustness. The next section presents results of calibration of the control-related parameters.

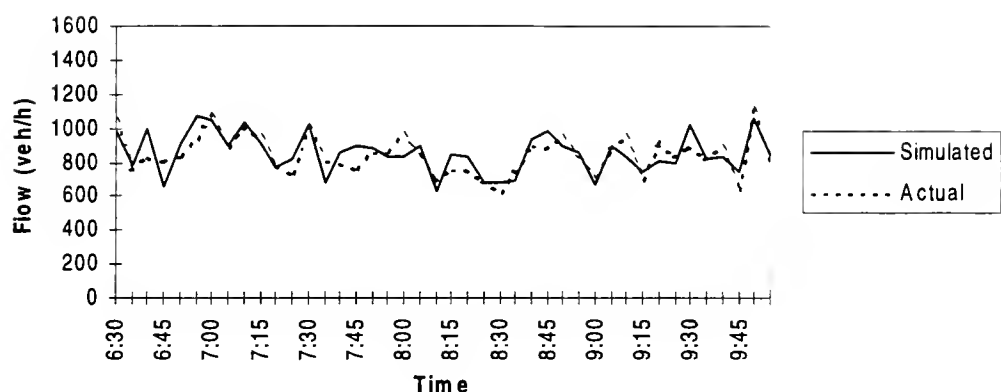


Figure 2.6.4 Flow Variation with Time at 1.25 km to Lane Drop in Continuous Lane

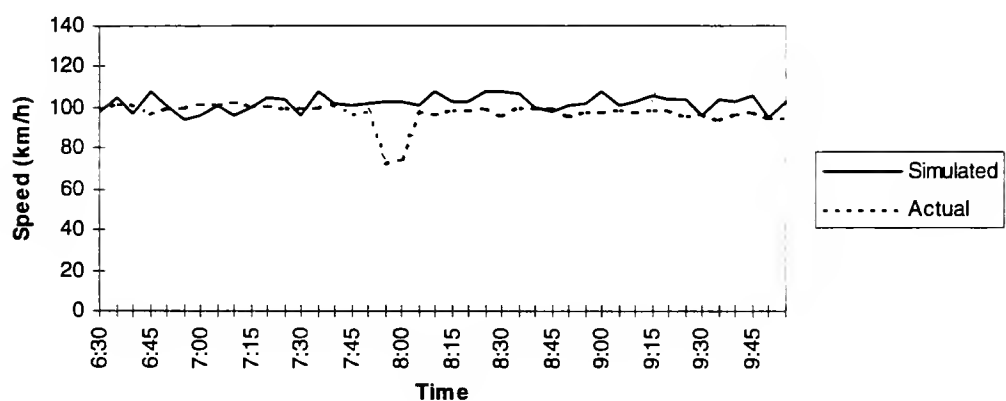


Figure 2.6.5 Speed Variation with Time at 1.25 km to Lane Drop in Continuous Lane

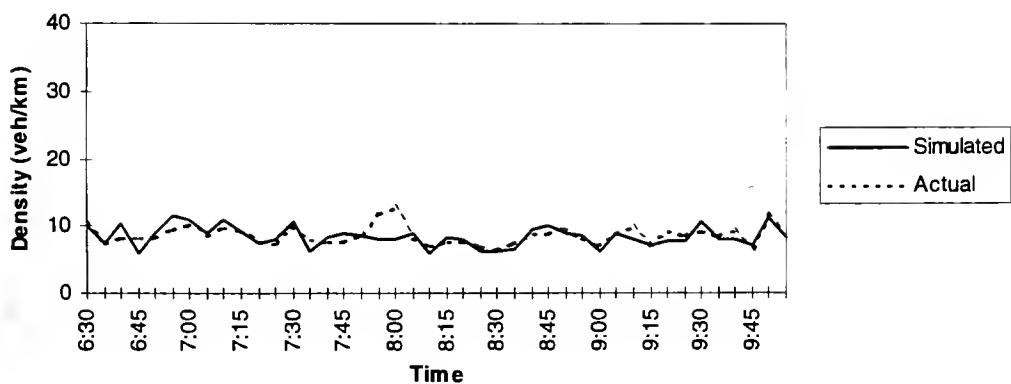


Figure 2.6.6 Density Variation with Time at 1.25 km to Lane Drop in Continuous Lane

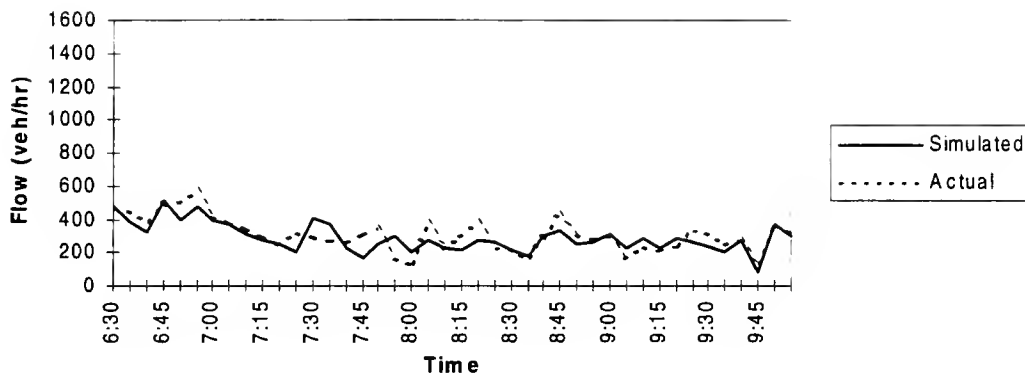


Figure 2.6.7 Flow Variation With Time at 1.25 km to Lane Drop in Dropped Lane

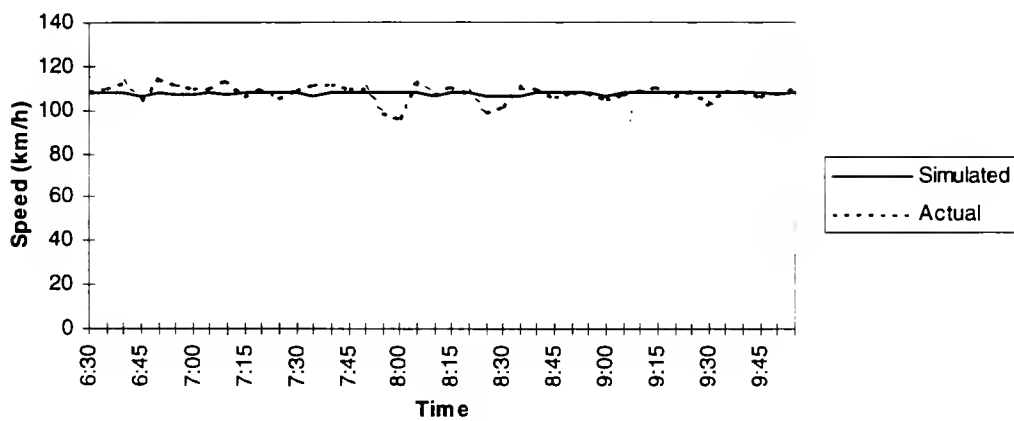


Figure 2.6.8 Speed Variation With Time at 1.25 km to Lane Drop in Dropped Lane

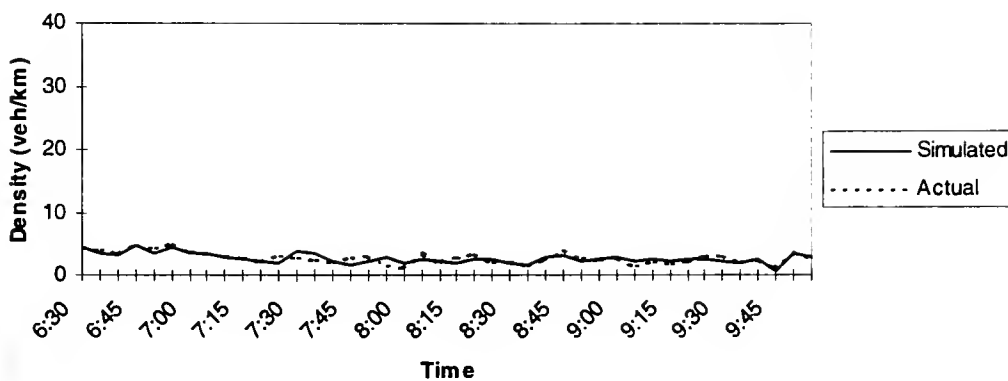


Figure 2.6.9 Density Variation With Time at 1.25 km to Lane Drop in Dropped Lane

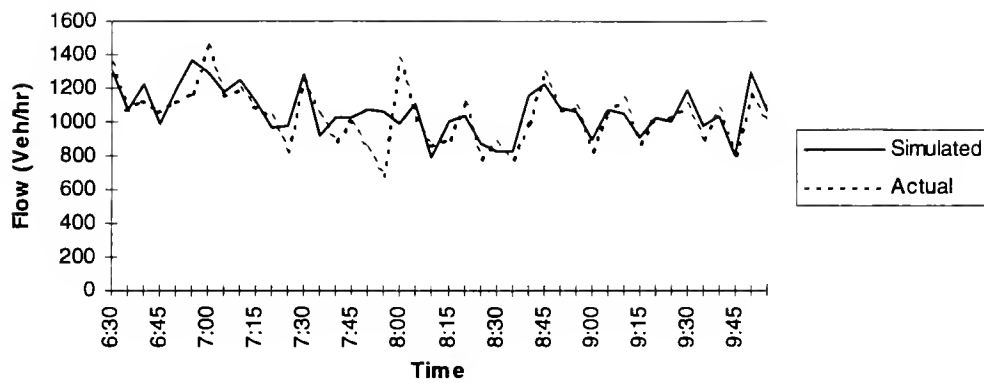


Figure 2.6.10 Flow Variation with Time at 0.05 km to Lane Drop in Continuous Lane

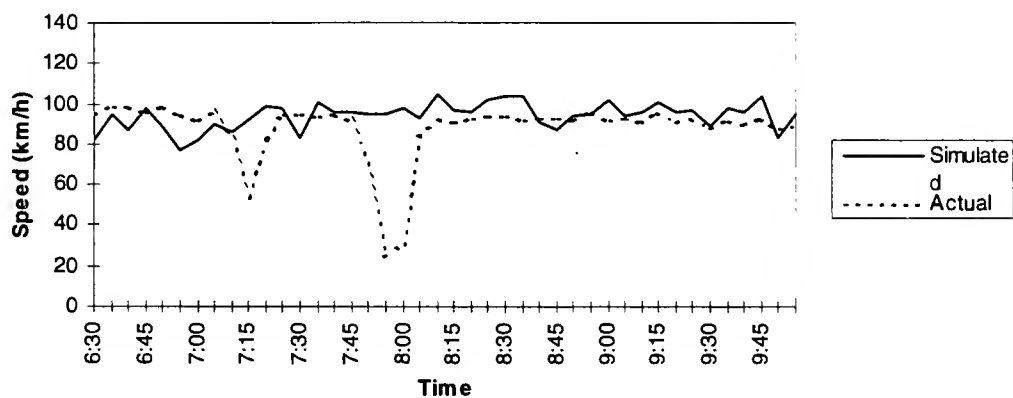


Figure 2.6.11 Flow Variation with Time at 0.05 km to Lane Drop in Continuous Lane

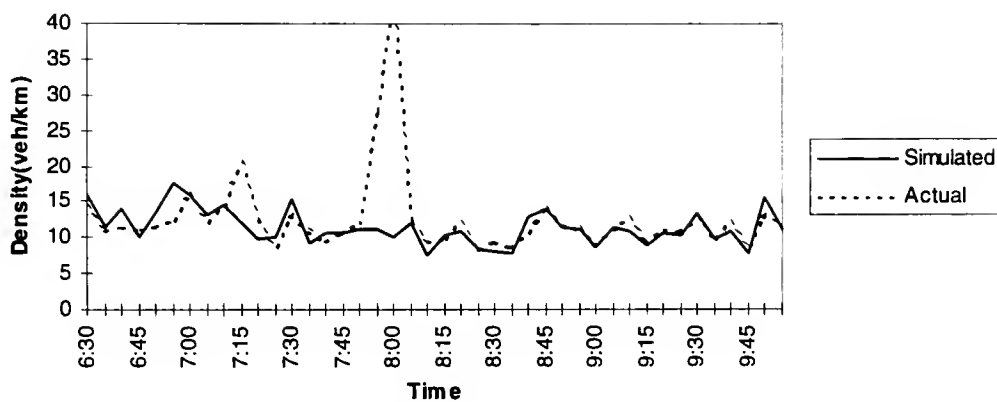


Figure 2.6.12 Density Variation with Time at 0.05 km to Lane Drop in Continuous Lane

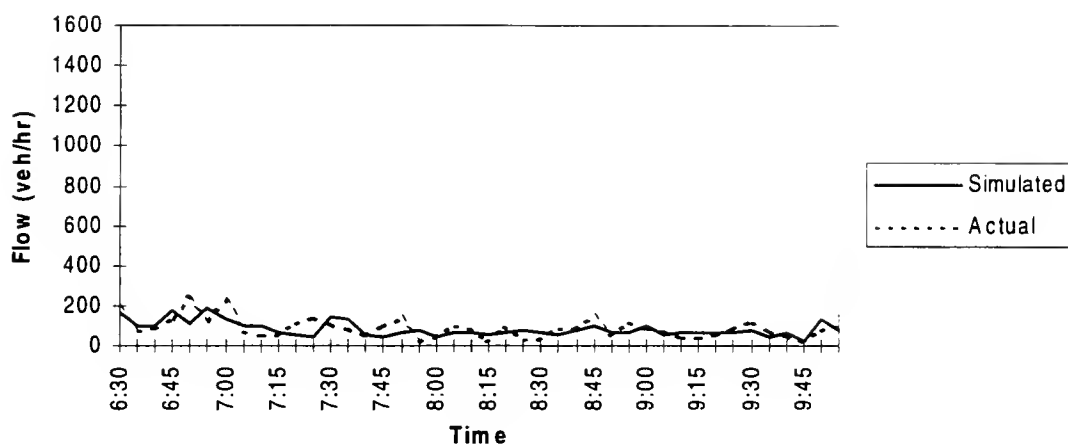


Figure 2.6.13 Flow Variation with Time at 0.05 km to Lane Drop in Dropped Lane

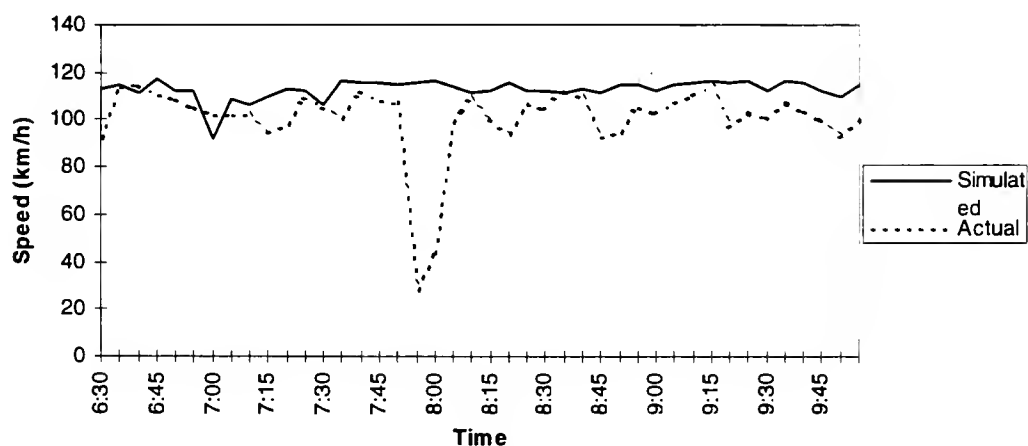


Figure 2.6.14 :Speed Variation with Time at 0.05 km to Lane Drop in Dropped Lane

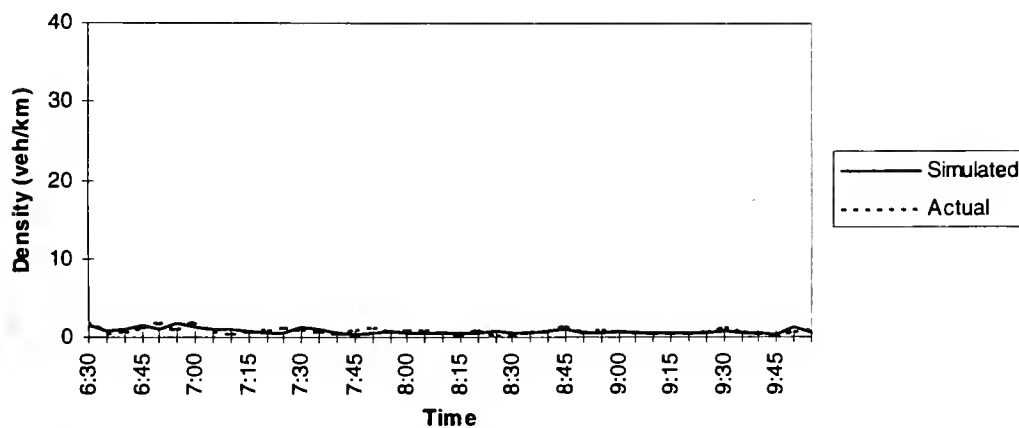


Figure 2.6.15 Density Variation with Time at 0.05 km to Lane Drop in Dropped Lane

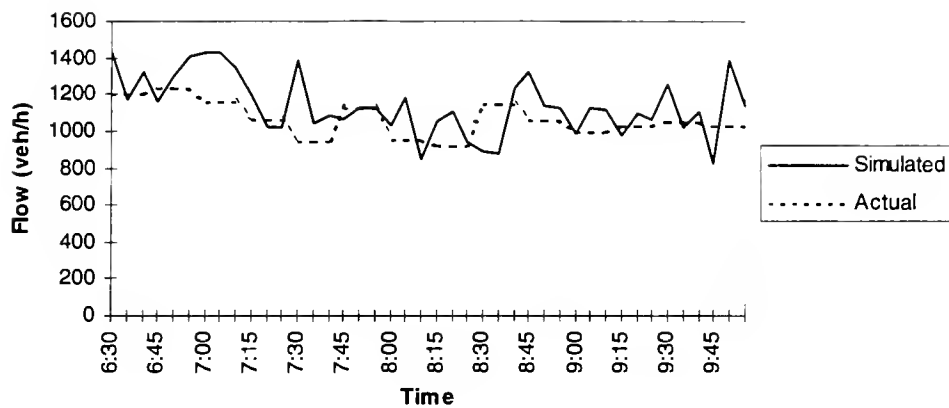


Figure 2.6.16 Flow Variation with Time at -0.80 km to Lane Drop in Continuous Lane

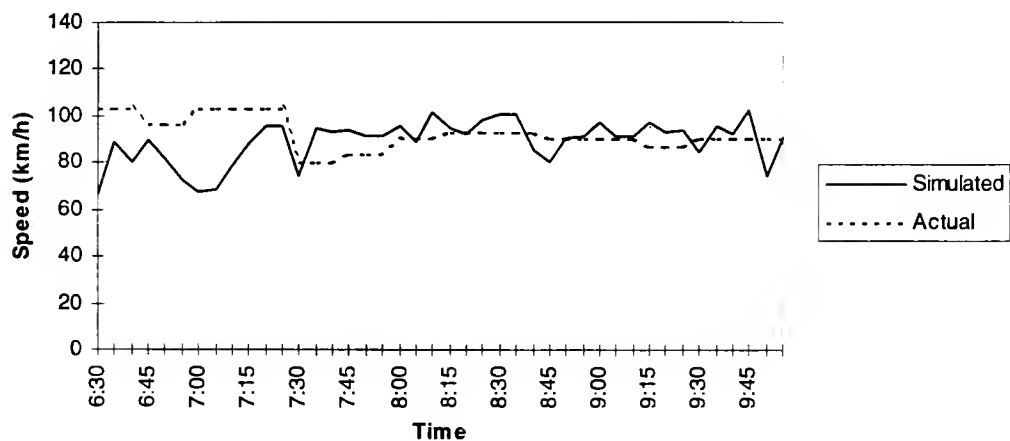


Figure 2.6.17 Speed Variation with Time at -0.80 km to Lane Drop in Continuous Lane

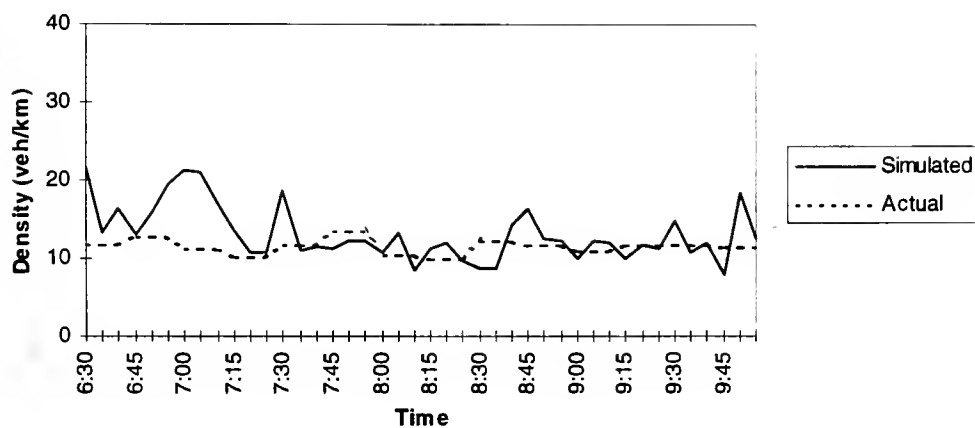


Figure 2.6.18 Density Variation with Time at -0.80 miles to Lane Drop in Continuous Lane

2.6.4 Third Stage of Calibration

The values of geometry factor, anticipation factor, viscosity factor, driver's perception factor, and risk factors have already been set and only the parameters representing the effect of the control boards remained not calibrated. We used the data collected at the same site, the I-69 southbound work zone, but this time with control boards in place. Table 2.6.3 gives the upper and lower bounds of the control-related parameter values used in the complex method. The Complex method was run number of times using different sets of randomly generated initial feasible points and the best set of values is reported in Table 2.6.4. The value of initial compliance rate, p_o^c is set to 0.2141, which states that 21% of drivers decide to change from the dropped to continuous lanes after passing the first active control board, in addition to the drivers changing lanes due the geometry effect. The value of incremental compliance rate α is set to 0.01405, which indicates that only about 1.5 % of the drivers decide to change lanes when they pass other active boards located downstream to the first one. This also indicates that drivers make the decision at the first active board whether or not to change lanes because of control. Very few drivers decide to change lanes after passing the first active board.

From Figures 2.6.19, 2.6.22, 2.6.25, 2.6.28, and 2.6.31, it can be seen that that the flow variation and distribution between lanes is captured very well. From Figure 2.6.31 it can be seen that the outflow rate is predicted reasonably close to the actual data. From Figures 2.6.21, 2.6.24, 2.6.27, 2.6.30, and 2.6.33, it can be seen that the trend in the variation of density is reasonably captured. Towards the end, there is a high density of vehicles observed in the continuous lane from Figures 2.6.21 and 2.6.27. This indicates that a breakdown has occurred very close to the lane drop point and propagated up to 1 km to the lane drop point. The model also predicts the breakdown at about the same time it is observed in the real data (Figure 2.6.27), but it is not propagated to a long distance. In the model, the system adjusted very soon and recovered from the breakdown. The novelty of the system could contribute to the breakdown. Drivers became more

cautious due to the presence of the new and unfamiliar control boards. Lower speeds and capacity are expected effects. Indeed, the speed reduction is seen from the Figures 2.6.20, 2.6.23, 2.6.26, and 2.6.29. However, the assumption in the model is that it predicts the long-term behavior of the drivers' compliance to control. Therefore, if the control had been placed one or two weeks before the data collection, the daily commuters probably would have become accustomed to the new control and exhibited behavior close to the predictions in the model.

Table 2.6.3 Control Parameters' Constraints

Parameter	Lower Bound	Upper Bound
Initial Compliance Rate	0.0000001	0.5
Incremental Compliance Rate	-0.2	0.2

Table 2.6.4 Calibrated Control Parameters

Parameter	Calibrated Value
Initial Compliance Rate	0.214
Incremental Compliance Rate	0.014

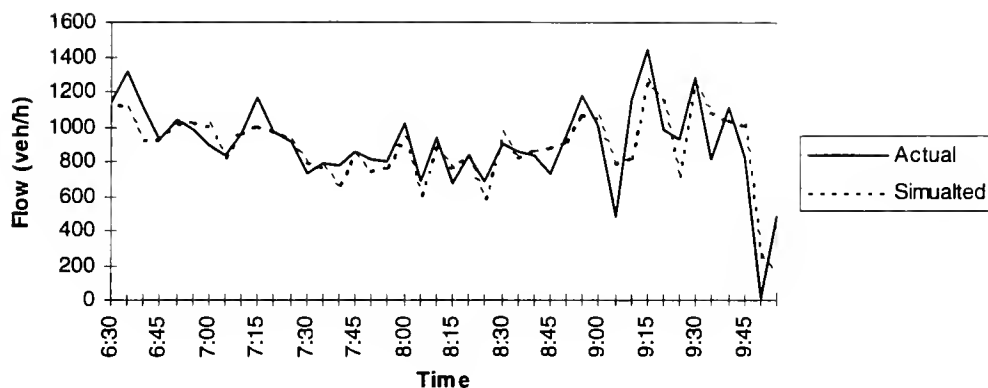


Figure 2.6.19 Flow Variation with Time at 1.25 km to Lane Drop in Continuous Lane

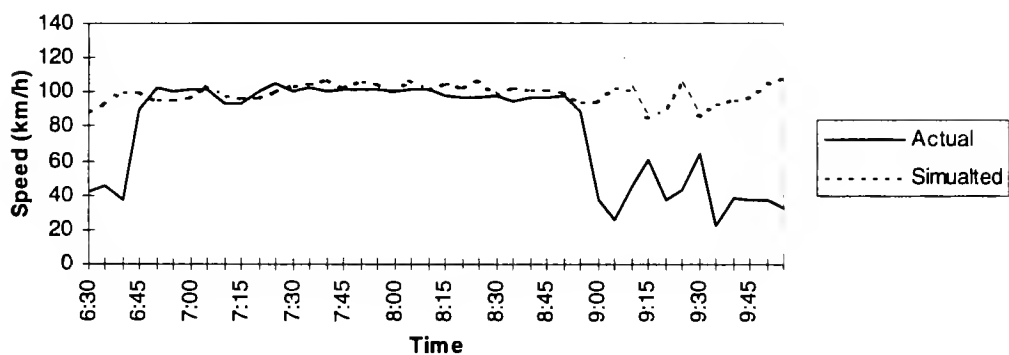


Figure 2.6.20 Speed Variation With Time at 1.25 km to Lane Drop in Continuous Lane

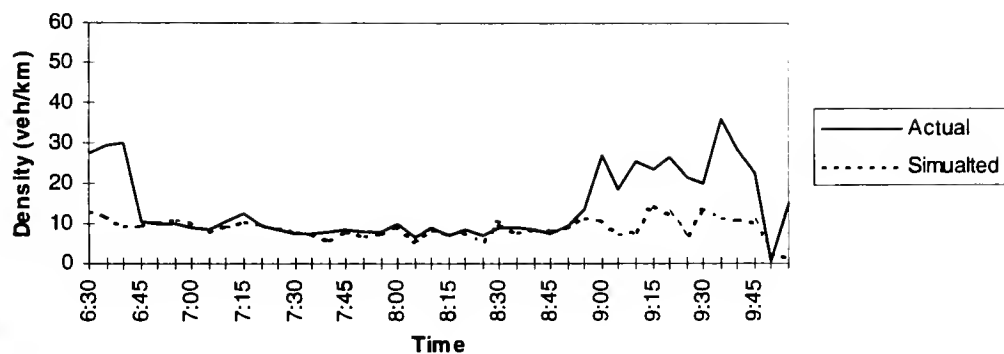


Figure 2.6.21 Density Variation With Time at 1.25 km to Lane Drop in Continuous Lane

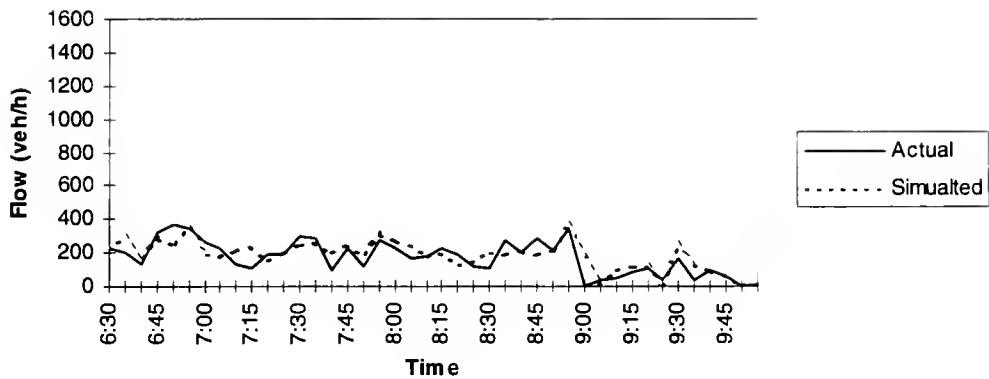


Figure 2.6.22 Flow Variation With Time at 1.25 km to Lane Drop in Dropped Lane

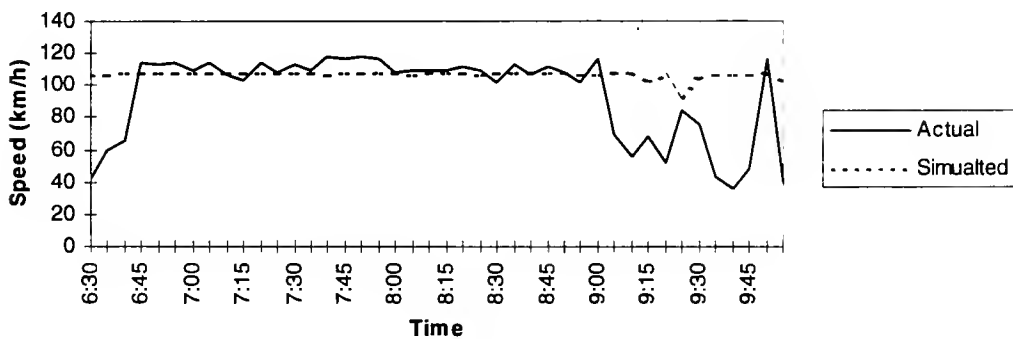


Figure 2.6.23 Speed Variation With Time at 1.25 km to Lane Drop in Dropped Lane

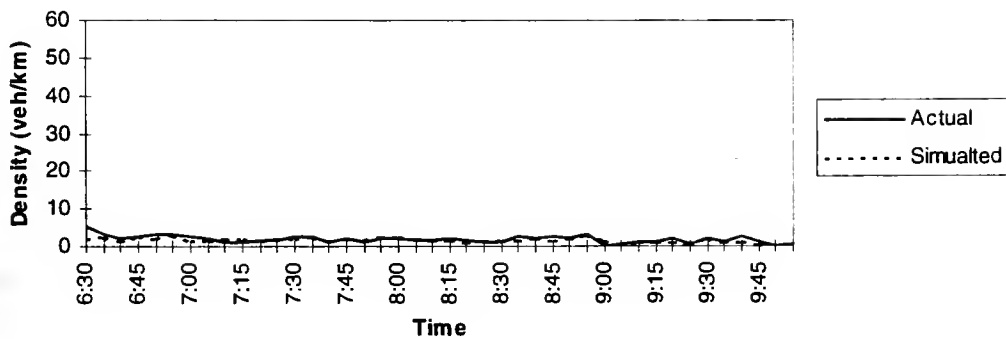


Figure 2.6.24 Density Variation With Time at 1.25 km to Lane Drop in Dropped Lane

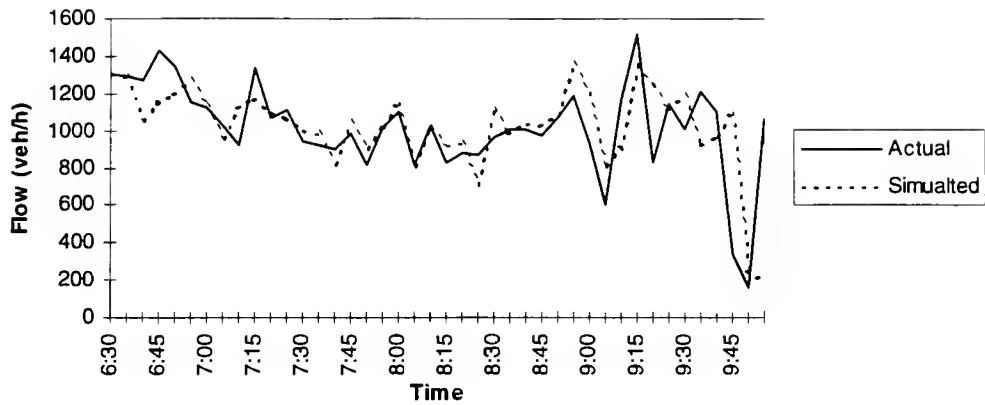


Figure 2.6.25 Flow Variation With Time at 0.05 km to Lane Drop in Continuous Lane

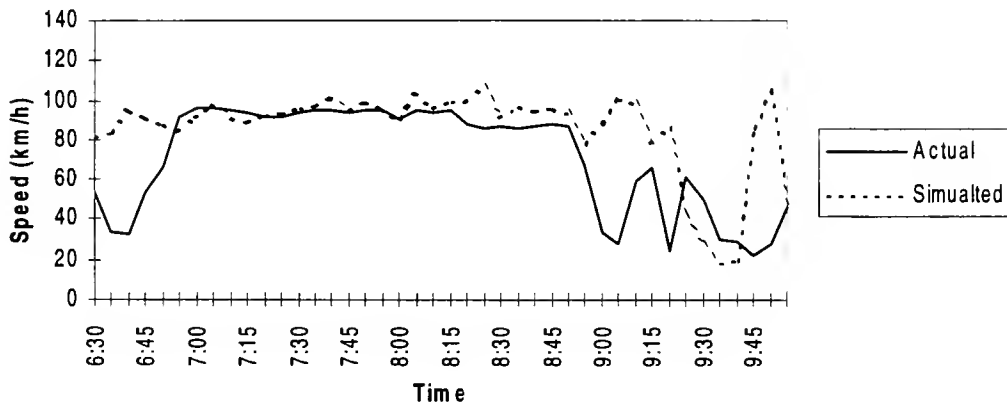


Figure 2.6.26 Speed Variation With Time at 0.05 km to Lane Drop in Continuous Lane

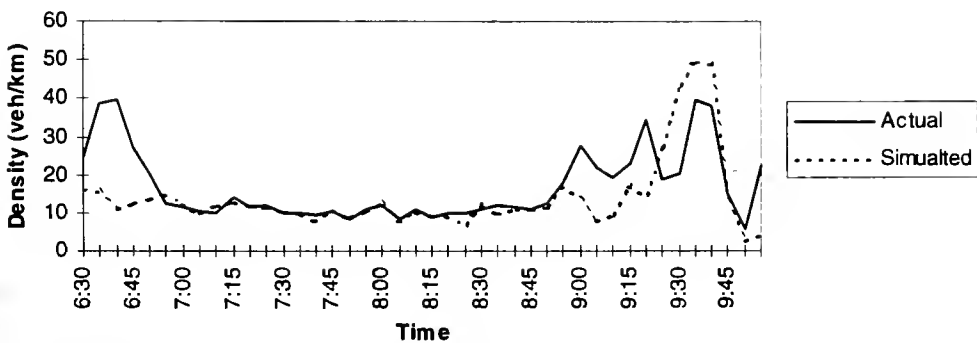


Figure 2.6.27 Density Variation With Time at 0.05 km to Lane Drop in Continuous Lane

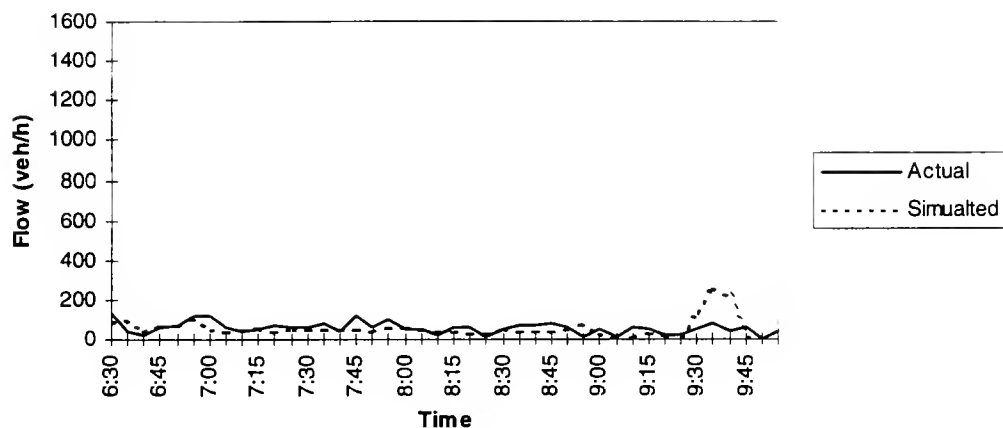


Figure 2.6.28 Flow Variation With Time at 0.05 km to Lane Drop in Dropped Lane

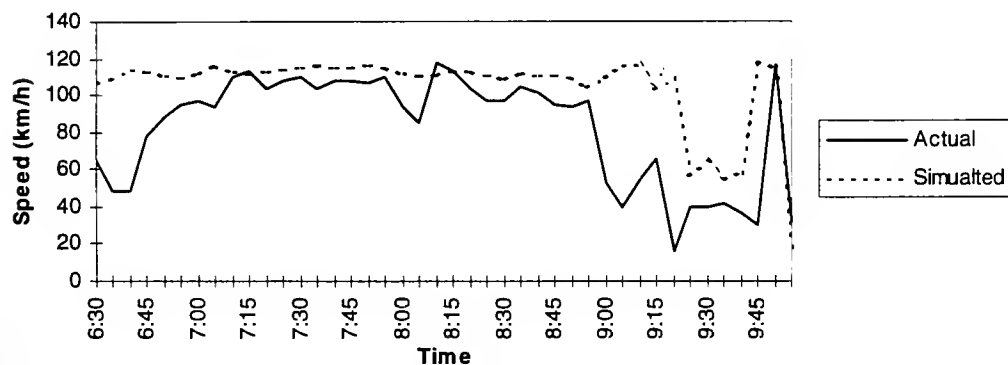


Figure 2.6.29 Speed Variation With Time at 0.05 km to Lane Drop in Dropped Lane

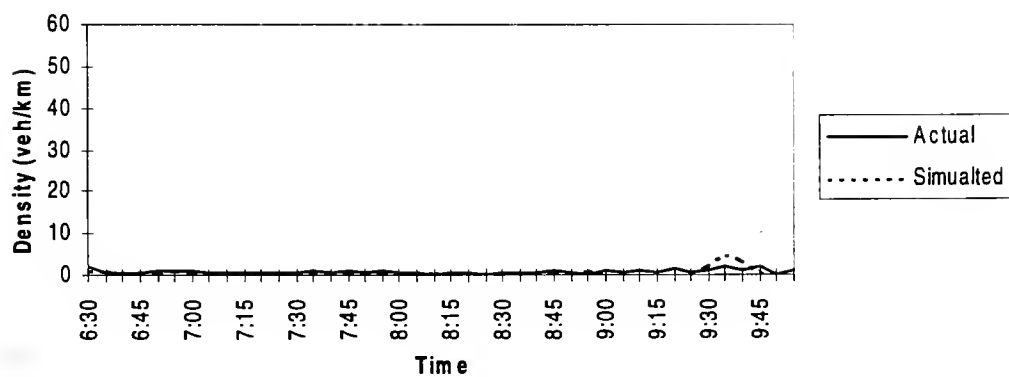


Figure 2.6.30 Density Variation With Time at 0.05 km to Lane Drop in Dropped Lane

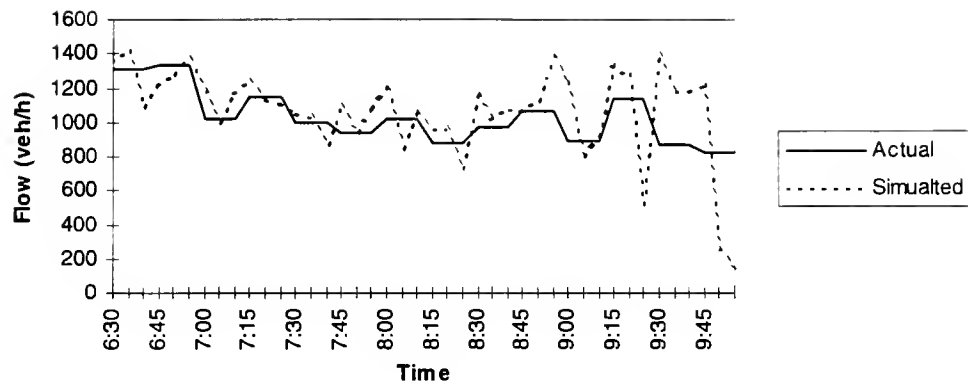


Figure 2.6.31 Flow Variation With Time at -0.80 km to Lane Drop in Continuous Lane

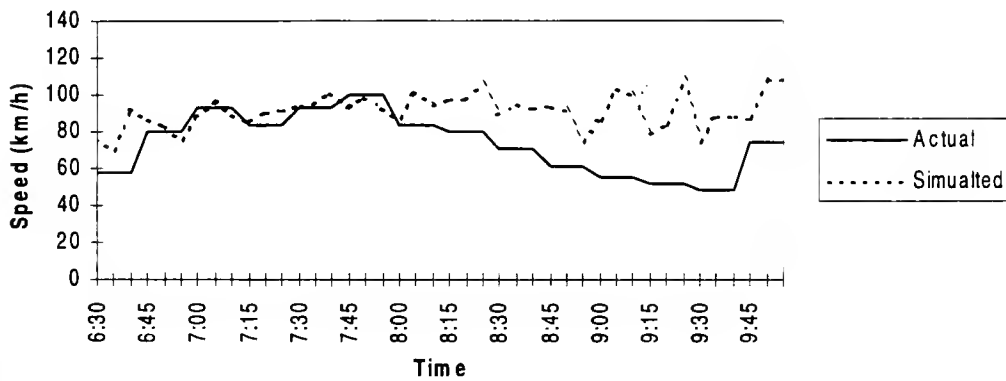


Figure 2.6.32 Speed Variation With Time at -0.80 km to Lane Drop in Continuous Lane

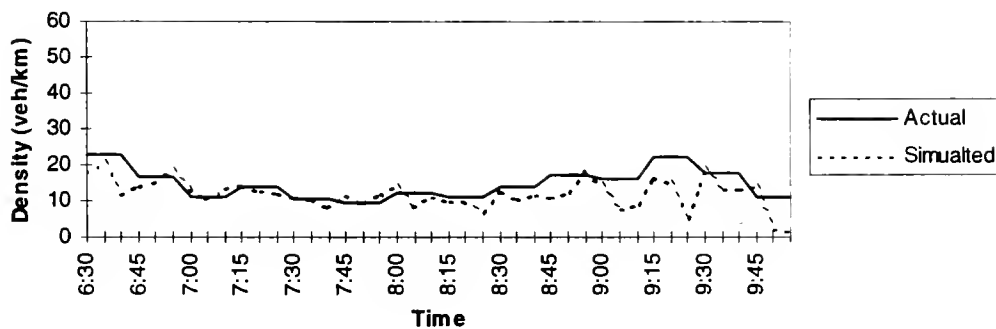


Figure 2.6.33 Density Variation With Time at -0.80 km to Lane Drop in Continuous Lane

2.7 EVALUATION OF THE MODEL

The results in Section 2.6 show that the model is capable of capturing the variations in traffic characteristics of the work zone for which the parameters have been calibrated. In order to test the validity of the calibrated model, the model must be checked as to whether it can capture the variations in traffic characteristics at other sites. This process is known as evaluation of the model. The model is evaluated using the data collected from the northbound merge in the I-69 work zone. This section presents the methodology and results obtained from the evaluation process.

2.7.1 Method

From Section 2.4.5, it is understood that the model requires input including flow and speed values for every lane i at the first section. These values have been provided in the data obtained from detectors at location No. 1 (Figures 2.5.3 & 2.5.4). Reasonable initial conditions were obtained by applying a warm-up time (about 10 minutes). The warm-up simulation starts with free-flow conditions, and density, speed, and flow are computed in every iteration of the warm-up time with all conditions set as discussed in Section 2.4.6.

The output from the model was aggregated and saved in the pre-specified time intervals and at the locations corresponding to the time intervals and the locations at which data were collected in the field. Then, graphs were plotted for traffic densities at these locations for both the actual and the simulated data to compare and evaluate the performance of the model.

2.7.2 Evaluation Results

This section presents the results in the evaluation process. As mentioned in the previous section, graphs were plotted for the actual and the simulated data for the northbound merge of the I-69 work zone. The model was first evaluated using the data collected without control boards in place. Figure 2.7.1 shows the variation of inflow rate in both the continuous and dropped lanes. This inflow was used as input to the model. We were informed by the INDOT personnel involved in the data collection that the traffic was partially blocked several times during the data collection time by the trucks engaged in the activities in the work zone. This could have caused the sudden changes in traffic characteristics. We did not have any precise description of the events we could use to replicate the real-world situation. Instead, we decided to check whether the model is able to produce a similar pattern if the continuous lane was partially blocked in the model for specific periods in a few sections downstream to the location 2. Figures 2.7.2 through 2.7.6 present the both the actual and the simulated densities with the partial blockage of the continuous lane.

Figure 2.7.2 and 2.7.3 show that the variations of density in the continuous and discontinued lanes at 1.25 km to the lane drop are very well captured.

Figures 2.7.4 and 2.7.5 show the variations of traffic density in both the actual and simulated cases at a distance of 0.05 km to the lane drop. This time, the traffic conditions in the model seem to be unstable more than in the real-world situation. Nevertheless, the overall trends in traffic changes over time and traffic distribution between lanes are replicated appropriately.

Figure 2.7.6 presents the variation in traffic density inside the work zone 0.80 km downstream of the lane drop. Again, the variations in the density are well captured. This indicates that the model replicates the real characteristics well.

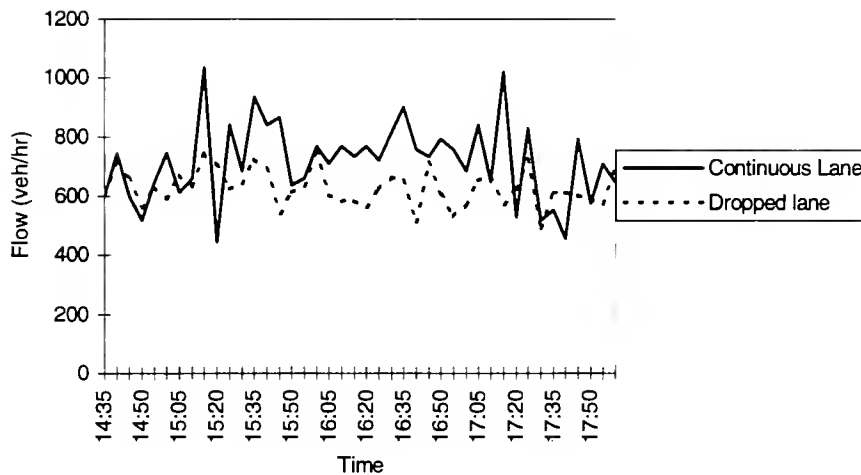


Figure 2.7.1 Flow Variation at Input Section Both in Continuous and Dropped Lanes

An attempt was made to adopt the same procedure to the with control case as was used for the without control case. Unfortunately, three detectors did not function properly during the data collection with control boards in place in the I-69 northbound work zone. The detector recording the inflow volume was one of those critical for the model evaluation. Moreover, we were informed that the traffic was blocked completely by workers in the construction zone for a few minutes three or four times during the data collection time. The model has not been evaluated for the case with control in the I-69 work zone. Unfortunately, in the I-74 westbound work zone, the control boards did not operate at all due to the absence of queue. Based on the indications provided by the calibration and the limited evaluation, we expect that the model should perform reasonably for the case with control.

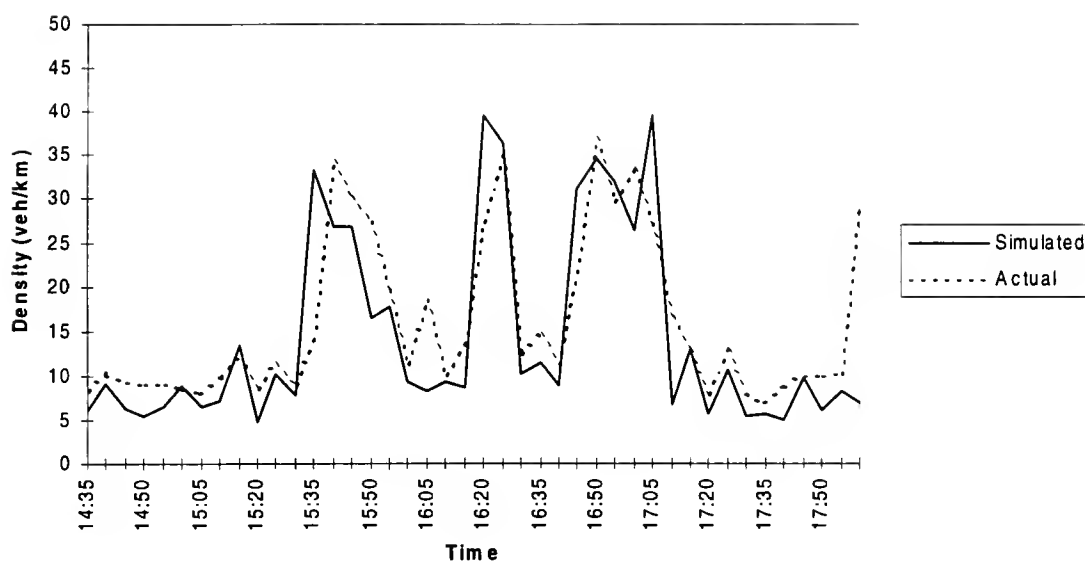


Figure 2.7.2 Density Variation With Time at 0.85 miles to lane drop in Continuous Lane

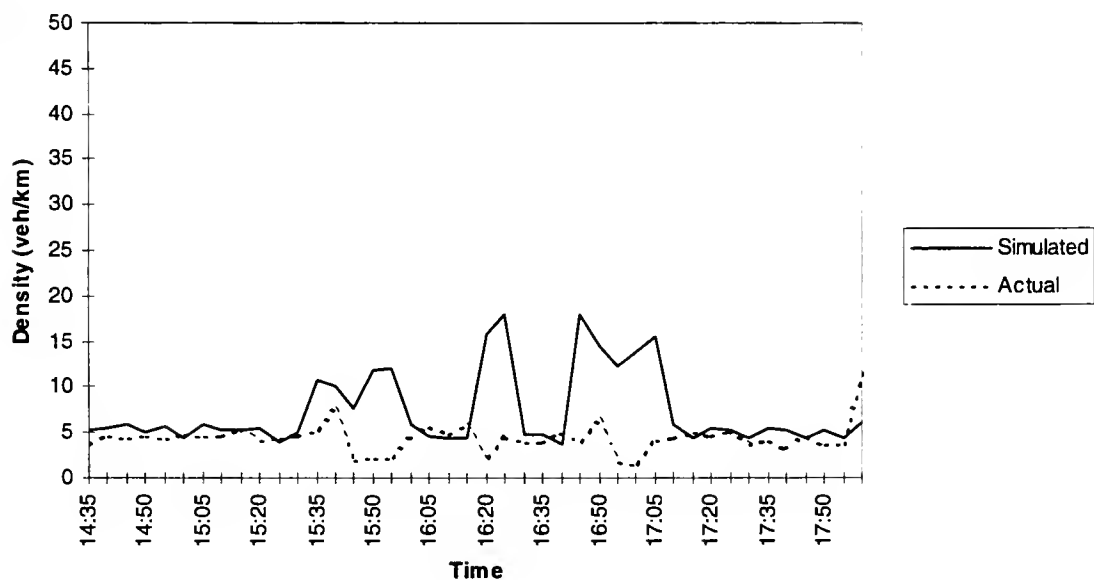


Figure 2.7.3 Density Variation With Time at 0.85 miles to lane drop in Dropped Lane

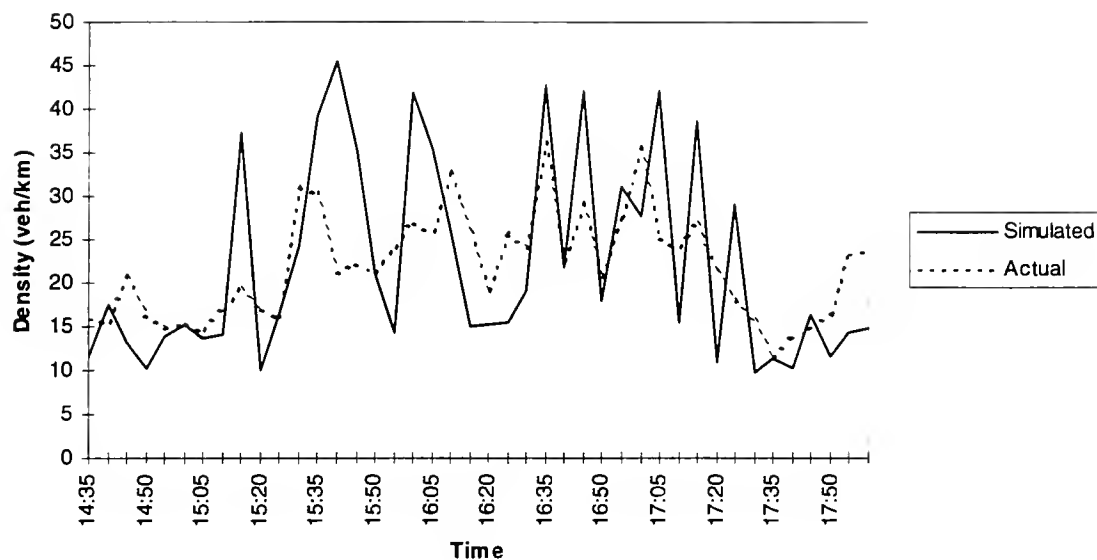


Figure 2.7.4 Density Variation With Time at 0.06 miles to lane drop in Continuous Lane

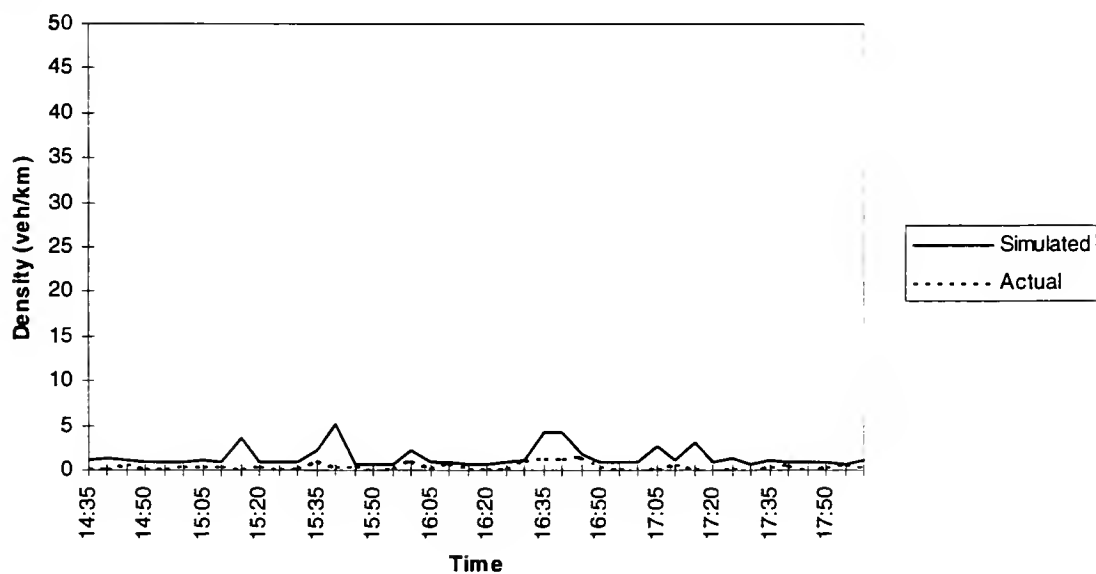


Figure 2.7.5 Density Variation With Time at 0.06 miles to lane drop in Dropped Lane

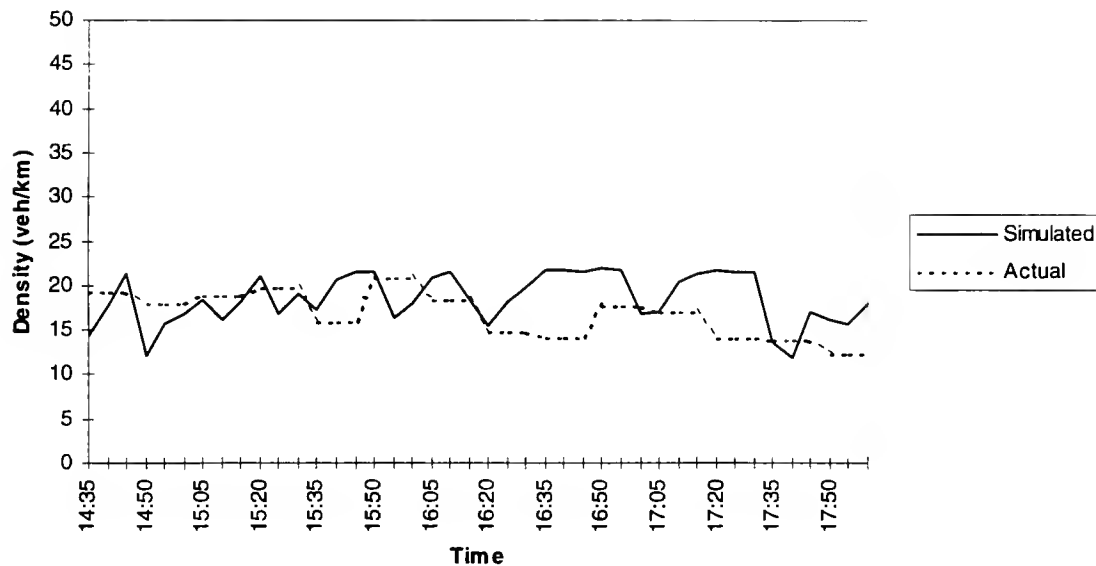


Figure 2.7.6 Density Variation With Time at -0.91 miles to lane drop in Continuous Lane

From the results obtained in the evaluation, it can be concluded that the model can capture traffic conditions satisfactorily even in the presence of partial temporary lane blockages.

2.8 SUMMARY AND CONCLUSIONS

A macroscopic modeling approach was selected to develop a fast traffic model. A continuum model has been extended for freeway traffic to incorporate drivers' response to the geometry, traffic conditions, and effect of traffic control. The incorporated factors are distance to the lane drop, opportunity for changing lane, and the presence of DO NOT PASS boards. All these factors were easily incorporated within the proposed theoretical framework of decision tree with the logit models.

The continuum model is discretized over both time and space to solve the partial differential equations numerically with complex boundary conditions. The values of density, speed, and flow are computed at a discrete point in space and time based on the previous time step values at respective locations.

The model has a number of parameters that needed to be calibrated using real data. For this purpose, real data was collected at three sites using magnetic detectors (speed and flow) and a distance-measuring vehicle (to measure distances of detectors from lane drop). The parameters were calibrated using a three-stage process. In the first stage, the speed-density relationship was calibrated as a three-regime model. The generalized equilibrium speed-density relationship requires only a free-flow speed for the site. In the second stage, the parameters that do not account for the control effect were calibrated. The total squared differences between the observed and simulated densities were chosen as an error (performance) function. A non-gradient search algorithm, complex method was used to calibrate these parameters. In the third stage, the parameters that account for the control effect were calibrated using the complex method.

The results of the model evaluation indicate the reasonableness of the model. The model predicts the variations in flow, speed, and density close to the reality when there are no interruptions in traffic. The model predicts the variation in traffic conditions satisfactorily when traffic interruptions are present, though the time and the magnitude of the capacity reduction must be known.

The presented version of the model describes the operation of a two-lane segment. The future effort is needed to add the following features: more lanes, exit and entrance ramps, and advance information about the work zone. The model in the current version is not able to handle multi-class traffic (passenger cars, trucks etc.). It can be partly accomplished by using an equivalent factor to convert heavy vehicles into cars. This factor can be calibrated as other model parameter.

The model with its extensions can be used to study the effectiveness of the control by comparing travel time near the work zone without control and with control. In addition, the control settings can be optimized so that travel time near the work zone is minimized based on the variation in travel time for different control settings.

The developed model and its extended version can be used in several ITS applications:

(1) Work zone traffic management with the dynamic traffic control using DO NOT PASS boards.

(2) Incident traffic management to predict travel times critical for the proper estimation of diverging vehicles from the freeway segment affected by the incident.

(3) Recurring congestion management with the on-line lane use control, speed control, and ramp metering.

2.9 LIST OF REFERENCES

Daganzo, C.F. (1994). The cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory; *Transportation Research - B*, vol. 28B, pp. 269-287

Dressler, R.F. (1949). Mathematical Solution of the Problem of Roll-Waves in Inclined Open Channels; *commun. Pure ans Applied Math.* 2, pp. 149-194.

Gazis, D.C., Herman, R., and Weiss, G.H. (1962). Density Oscillation between Lanes of Multiple Lane Highway, *Operations Research*, 10, pp. 658-667.

Hauer, E., and Hurdle, V.F. (1979). FREFLO: Discussion of the Freeway Traffic Model. *Transportation research Record No. 722*, pp. 75-77.

Lighthill, M.J. and Whitham, G.B. (1955a). On Kinematic Waves I. Flood movement on long rivers. *Proc. Roy. Soc. London* 229, pp. 281-316.

Lighthill, M.J. and Whitham, G.B. (1955b). On Kinematic Waves II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London* 229, pp. 317-345.

May, A.D. (1990). *Traffic Flow Fundamentals*. Prentice Hall, NJ, pp. 301-303.

Michalopoulos, P.G., Kwon, E. and Khang, E.G. (1991). Enhancements and Field Testing of Dynamic Simulation Program. *Transportation Research Record*, #1320, pp. 203-215.

Michalopoulos, P.G. (1988). Analysis of Traffic Flow at Complex Congested Arterials. *Transportation Research Record*, #1194, pp. 77-86.

Payne, H.J. (1979). FREFLO: A macroscopic Simulation Model for Freeway Traffic. *Transportation Research Record* 772, pp. 68-77.

Prigogine, I., and Herman, R. (1971) *Kinematic Theory of Vehicular Traffic*, Am. Elsevier Publ., New York.

Stephanopoulos, G. and Michalopoulos, P.G. (1981). An Application of Shock Wave Theory to Traffic Signal Controls. *Transportation Research Journal*, Vol. Q5b, pp. 35-51.

Transportation Research Board, (1975). *Traffic Flow Theory*. Special Report 165,. Gerlough, D.L., and Huber, M.J. (editors).

Transportation Research Board, (1996). *Traffic Flow Theory*. Special Report 165,. Garther, N.H., Messer, C.J., and Rathi, A.K. (editors), to be published.

Whitham, G.B. (1974). *Linear and Nonlinear Waves*. Wiley & Sons, NY, pp. 68-95.

CHAPTER 3 SIMULATION USER'S MANUAL

3.1 INTRODUCTION

The simulation model is a macroscopic model. It has been coded in C++. The model is capable of simulating the conditions of traffic on a freeway. The user has to specify the freeway characteristics, control characteristics, periodic inflow data in each lane, and input period in separate input files, which will be described later in this manual. The program generates many output files, which will be described later in this manual. There are two sections in the manual. The first section describes the input file formats and the second section describes the output files and the output data.

3.2 INPUT FILES

The model requires six input files and each file is described with its content and format in this section. The input files are

D_Chars.dat

F_Chars.dat

C_Chars.dat

O_Chars.dat

lane1.inp

lane2.inp

Each input file is set for a specific group of parameters, which will be described later in this section. For example, D_Chars.dat is set for specifying the parameters corresponding to driver's characteristics, F_Chars.dat is set to specify the freeway characteristics, etc. The order of the variables can be anything in these files, but all the files require all the values to be specified. The names of the parameters must be the same as shown and must be followed by an '=' and then the value of the parameter is specified. The blank spaces

before the parameter name, '=' and the values can be any number. The following paragraphs describe the variables specified in each file.

3.2.1 D_Chars.dat

This file contains the input values for variables that describe the drivers' behavior. These values have been calibrated and so the user need not change these values. This file contains the values in the following format.

```
a = 0.0020041
g = 0.0242047
T1 = 3.0
tau1 = 5.0
C0 = 1.4342
beta = 6691.29
delta = 0.0119894
Viscosity_Coeff = 25.33
p_c = 0.2408715
p_c_alpha = 0.0917462
```

The parameter values are calibrated for the first-order continuum models.

Here the parameter a corresponds to the geometry parameter. The value of a reflects the effect of geometry on the driver's behavior. The proportion of drivers who comply with control at any distance is given by $e^{-a \cdot l}$ where l is the distance to the lane drop and expressed in meters.

Value g is the parameter that accounts for the reduction in the critical space gap and the critical time gap as the lane merge point is approached. The critical gap is

assumed to decrease exponentially from a *MAX* value to a *MIN* value with distance to the lane drop. The following is the equation used. Distance l is expressed in meters.

$$critical_gap = MIN_VAL + (1 - e^{-g \cdot l}) \cdot MAX_VAL$$

Tl and tau_l are the relaxation time and the fixed time for lane change maneuver. The relaxation time is the time a driver needs to adjust his current speed to the desirable speed. This parameter appears in the second-order continuum model. The fixed time for lane change maneuver is the time measured between the instance when the driver found sufficient gap for lane change and the instance when the driver completed the lane change maneuver. The value of tau_l varies from four to six seconds. These are expressed in seconds.

CO is the anticipation coefficient. This value reflects the driver's response to the density changes down stream to their position, i.e., the driver anticipates the conditions downstream and reacts accordingly. If the down stream density is higher than the current density, then the anticipation term makes the driver decelerate and vice versa. This is used in the traffic propagation term and in the traffic propagation equation of the second-order continuum model.

Parameters β and δ are the drivers' perception and risk factors. The parameters are used in calculating the likelihood that a driver is willing to change lanes based on traffic conditions.

The parameter δ can be interpreted as the risk associated with the lane change, expressed in the equivalent density. Since this requires an attenuation effect, likelihood δ is added to the density in the other lane and not to the density in the lane currently used by the driver. This means that the higher the value of δ , the smaller the number of drivers willing to change lanes. If δ is zero and the density in both lanes is the same, then the likelihood of choosing a lane is $\frac{1}{2}$, i.e., both lanes are equal choices. If the parameter β is

zero, irrespective of density and risk factor, the likelihood of choosing a lane is always $\frac{1}{2}$. The higher the value of β , the lesser will be the likelihood of a driver willing to change to the higher density lane.

Parameters p_c and p_{c_alpha} are the parameters that correspond to the effect of control. p_c represents the first board compliance rate, and p_{c_alpha} represents the incremental compliance rate. In other words, p_c states the likelihood a driver would comply to the first active control board he passes and p_{c_alpha} states the increase in the likelihood when he passes an additional active control board. For example, $p_c = 0.2408715$ indicates that 24 percent of drivers would comply to the first board in which specified by the user they encounter, and $p_{c_alpha} = 0.0917462$ indicates that for every next board there will be a nine percent increase in the compliance rate.

All of these values are calibrated for the real data when traffic characteristics are expressed in SI system of units (distance in meters and time in seconds).

3.2.2 F Chars.dat

This file contains the input values for describing the freeway characteristics and the simulation time. These values must be specified by the user based on the freeway geometry and the accuracy of the results desired. This file contains the values in the following format.

```
Freeway_length = 10
Uf = 105.00
No_of_lanes = 2
DROPPED_lane = 1
Sim_Time = 24
Segment_Length = 50
deltat = 1.0
```


The first parameter *Freeway_length* is used to specify the length of the freeway section to be simulated. This is expressed in kilometers. The freeway length is measured between the inflow point and the lane drop point. *Uf* is the variable used to specify the free flow speed in that freeway section. This value is used to calculate the speed-density relationship for the freeway section near the work zone. This is expressed in kmph. The lanes are numbered 0 and 1, with 0 indicating the driving lane and 1 indicating the passing lane. The dropped lane is specified using the variable *DROPPED_lane*. This value is important in identifying the input file for inflow data in the dropped lane. For example, if *DROPPED_lane* =1, then *lane2.inp* is the inflow data for the dropped lane.

(Note: The index in the file name is one more than the number specified for the *dropped_lane* variable. This is done because the index for arrays start from zero in C++.)

Sim_Time is the time for which the simulation is run. This value is expressed in hours. *Segment_Length* is the length of each segment used when the freeway is discretized into small sections. This value is expressed in meters. The last value *deltat* is the incremental time step in the calculations. This is expressed in seconds and is usually on the order of 0.5 to 2 seconds. The smaller the values of *deltat* and *Segment_Length*, the more accurate the results and more computation time.

3.2.3 C Chars.dat

This file contains the input values for describing the control characteristics, i.e, the number of DO NOT PASS boards, threshold value for activation, aggregation time, detector length, etc. These values must be specified by the user based on the desired control settings. This file contains the values in the following format.

No_of_Boards = 3

Location	Thrld_ocpnecy	Detector_Ingt	Impact_dst	Min_active-time
150	0	1.8	100	5
300	10	1.8	100	5
450	15	1.8	100	5

Aggregation_period = 5

Vehicle_length = 6.0

The first value in this file corresponds to the number of DO NOT PASS boards placed. This is followed by a line with header for specifying the control characteristics for control boards. This line reads as 'Location Thrld_ocpnecy Detector_Ingt Impact_dst in_active-time'. The order of these columns is important because the values are read from this file in the same order. Each row corresponds to one detector specifying characteristics such as threshold occupancy, detector length, impact distance, and minimum activation time. The number of rows specified here must match the number of boards specified. *Location* represents the distance of the control board from the lane drop, expressed in meters. *Thrld_ocpnecy* represents the threshold value of occupancy for activating this board, expressed as a percent value. *Detector_Ingt* is the length of the detector used in the field to activate this control board, expressed in meters. This value is used in computing the occupancy of the detector based on the density of traffic at its location. *Impact_dst* is the variable that determines the distance from which the drivers respond to the first active control board, expressed in meters. *Min_active_time* is the minimum time required between the turning on and turning off the boards, expressed in minutes.

The next value is *Aggregation_period*, specified in minutes. This value must be provided by the user to reflect the aggregation period set in the detectors and helps in determining the occupancy of the detector for the specified aggregation period. *Vehicle_length* is the average length of the vehicle, expressed in meters.

3.2.4 O Chars.dat

This file contains four values for specifying the inflow data interval and specifying the intervals for collecting the output statistics. The format of this file is as follows.

```
Inflow_interval = 60
Normal_report_interval = 30
Detail_report_interval = 60
Detail_switch=9
```

Inflow_interval is the variable used to specify the inflow time interval value. If the value for this variable is 60, it means that the files *lane1.inp* and *lane2.inp* has inflow rate and speed for every one-hour period. *Normal_report_interval* is used to specify the time interval for collecting the important output statistics like continuous lane travel time, average periodic travel time, and the number of active control boards. *Detail_report_switch* is used to specify the time interval for getting the density, speed, and flow rate profiles at a given instant over the entire length of the freeway segment. All of these values are expressed in minutes. As this takes too much memory space and is not usually required, the output can be controlled using the variable *Detail_switch*. If *Detail_switch* = 1, then the detail report is generated. If *Detail_switch* = 0, then the detail report is not generated.

3.2.5 lane1.inp & lane2.inp

The first line in these two files are header lines specifying flow and speed. So, the first column always corresponds to flow rate and the second column corresponds to speed. The flow rate is expressed in veh/hr and speed is expressed in km/hr.

3.3 OUTPUT FILES

The program generates four output files for different values. The list of output files generated by the program is as follows.

- detail.out
- overall.out
- report.out
- C_ln_tme.out

3.3.1 detail.out

This file has the variation of flow, speed, and density over the entire freeway length for various instants of time. These instances are separated by the time specified by the value of *Detail_interval* in the O_Chars.dat file. The output is arranged in four columns. The first column is for distance to the lane drop expressed in meters, the second column is for density of traffic expressed in veh/km, and the third column is for speed expressed in km/hr.

3.3.2 overall.out

This file reports the total number of vehicles initially present in the system before the simulation is started. This file reports the number of vehicles that entered the system during the simulation period and the total vehicle hours spent in the system. The average travel time per vehicle for the entire simulation period is reported as the ratio of these two values. Additionally, the number of vehicles leaving the system during the entire simulation period is reported.

3.3.3 C ln tme.out

This file reports the continuous lane travel time for drivers starting from the entry point to the system separated by time gap specified by *Normal_report_interval* in the file *O_Chars.dat*. Continuous lane travel time is the travel time experienced by a driver if the driver travels in the continuous lane from the point of entry up to the exit (outflow) point without shifting to the discontinuous lane at any point. The continuous lane travel time is expressed in minutes in this file. This travel time can be used as a measure to evaluate the effect of control on the drivers' behavior in reducing the delay for drivers sticking to the continuous lane.

3.3.4 report.out

This file contains the average travel time for drivers entering the system for the time intervals specified by *Normal_report_interval* in the file *O_Chars.dat* and the number of control boards that are active at the beginning of the time interval. This file has three columns. The first column represents the start time of the interval expressed in minutes, and the second column represents the average travel time for all the users entering the system during that period starting from the time specified. These travel time values are expressed in minutes. The third column gives the number of active control boards at the beginning of the time interval.

CHAPTER 4 SIMULATION EXPERIMENT

4.1 OBJECTIVES OF THE SIMULATION

The simulation experiments had three primary objectives.

1. Evaluate the ILMS impact on the traffic performance on approaches to freeway work zones.
2. Develop the warrants for the use of ILMS.
3. Develop a set of practical rules of the ILMS setting.

These objectives have been accomplished in the following steps.

1. Determine the simulation cases that sufficiently cover the range of traffic conditions on four-lane rural freeways.
2. Select the Measure of Effectiveness (MOE) to evaluate the quality of system's performance.
3. Find the optimal settings for each simulated case using the selected MOE.
4. Develop the rules that generate the settings of ILMS that are as close as possible to those obtained from the optimization.
5. Simulate the cases of freeway merge without the ILMS in place. Compare the traffic performance of the freeway merge without and with the ILMS.
6. Determine conditions that justify the use of the ILMS (warrants).

The first-order version of the continuum model developed in this project and described in Chapter 2 has been used to conduct the simulation experiments.

4.2 SIMULATED CASES

It is important that simulated cases represent a sufficient spectrum of traffic conditions that allow for investigating the trends in the results and for formulating the rules of ILMS setting. The most important factor of the system setting is the length of the congested

segment and its development over time. To ensure a significant variation of the length of the congested segment, simulated cases have been defined through the variation of the overflow volume. A typical volume profile assumed for the simulation is shown in Figure 4.1.

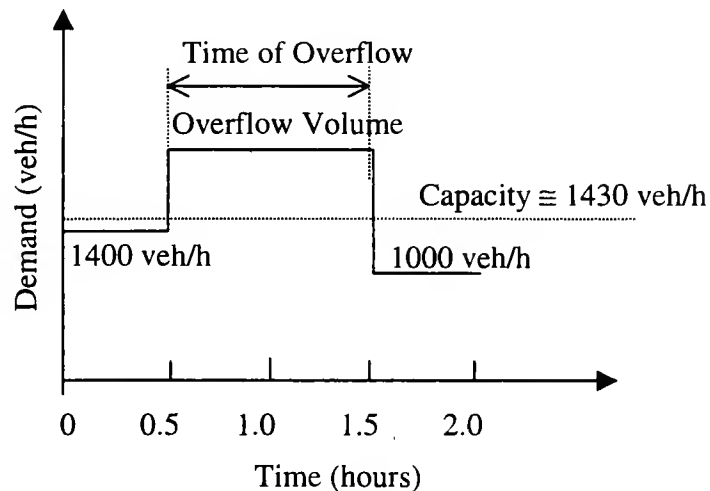


Figure 4.1 Illustration of Traffic Demand and Work Zone Capacity

Four traffic demands (called overflow volumes) were simulated for a two-lane rural freeway: 1500, 1600, 1700, and 1800 veh/h.

4.3 MEASURE OF EFFECTIVENESS

Three main criteria can be used to evaluate the traffic performance under the presence of IMLS:

1. The impeding effect of the work zone on the motion of vehicles, which can be represented through the average travel time or delay along the approach to the work zone,
2. The level of safety measured through the expected crash frequency on the merge to the work zone, and

3. The users' perception of the new system, which can be linked with the distribution of delay between the drivers who comply with the signs and the drivers who don't.

The field observations described in Chapter 2 indicate that a significant increase in capacity caused by the ILMS should not be expected. In fact, a slight reduction in capacity was observed at two sites immediately after the prototype system was applied. This effect could stem from the surprise the drivers experienced when seeing the new system for the first time. The caution of drivers, manifested by lower speeds and longer time headways, could be the primary reason for the reduced capacities.

The simulation model mimics the capacity effect of ILMS, which indicates the soundness of the model calibration, but it also raises the same concern as for the empirical results. Recognizing the above limitations, the authors decided to enforce in the simulation the same capacity before and after the IMLS is deployed. The first-order version of the model has been re-calibrated and used in the simulation experiments. Since the capacity and volume characteristics are only factors of delays, the assumed lack of the capacity impact implies the lack of the ILMS impact on the average delays and travel times. Thus, this criterion was not used to optimize the system settings.

The accident frequency is another important MOE that was considered. Unfortunately, no ILMS had been used prior to our research; thus, no crash records were available to analyze the safety impact. Even the traffic conflict technique requires the ILMS in place for a sufficient period of time to avoid biased results of the first few days of the system operation. Instead, an indirect measure -- average deceleration -- is used, which regards the fact that the deceleration maneuver is a response to the risk of crash (decreasing spacing between vehicles) while the acceleration maneuver is not (growing spacing between vehicles).

The last criterion is the perception of the system by the drivers. Since the majority of drivers comply with the system and stay in the continuous lane, any reduction in their delay must improve the overall perception of the ILMS. Of course, this benefit is associated with the increase in delays of the drivers who ordinarily would continue using the discontinuous lane until the latest possible moment. We have decided to use the average travel time in the continuous lane as the MOE for this project. The ILMS is optimized to minimize the delay of the drivers complying with the signs.

The second MOE (related to drivers' perception) is the change in the number of vehicles that pass the drivers in the continuous lane. This value has been calculated in the following way. First, an average number of departures from the freeway segments during travel in the continuous lane is estimated using two model outputs: total number of departures during the simulation period and average travel time in the continuous lane:

Average no. of departures =

Total number of departures x Average travel time in continuous lane / Simulation time.

The average number of departures in the continuous lane has been calculated for two cases: without ILMS and with ILMS. The difference between the numbers of departures in these two cases equals the change in number of passing maneuvers observed by the driver who uses the continuous lane.

4.4 OPTIMIZATION OF INDIVIDUAL PARAMETERS

For all of the cases under consideration, the system set-up is optimized for the following parameters (optimization variables):

1. Number of dynamic signs,
2. The total distance taken by the dynamic sign layout,
3. The spacing of the signs, and
4. The threshold detector's occupancy.

The aggregation period of detector output and the minimum activation time for the sign are important for the system stable operation. It is rather obvious that if the traffic conditions were changing in a smooth and continuous manner without any random jumps in traffic density, then these two settings should be short to make the system "snappy". However, due to the randomness in traffic characteristics, an aggregation period and a minimum activation time that are too short would cause frequent activation and deactivations of the signs even when the work zone does not experience permanent congestion. Since this behavior of the system is easy to observe in the field and is caused by the factors that are not represented in the simulation model, proper values of the two settings have been established by INDOT personnel based on their field experience with the prototype system made available in the later phase of the project.

All of the settings, except the third, have a clear meaning. The spacing of the signs needs more discussion. As shown in Figure 1.1, the first three signs are static with fixed distances between them. They are recommended to facilitate the law enforcement ramification of the system. This issue is further described in one of the next paragraphs. The remaining signs are dynamic and their position is subject to optimization. In the most flexible formulation of the optimization problem, the individual distances between any two consecutive signs are represented by an independent variable, which would add $n-1$ new variables (n = number of dynamic signs) to the optimization procedure. To reduce the number of variables, we propose to spread the dynamic signs along the freeway using the following recursive formula:

$$x_i = x_{i-1} \cdot k \quad (4.1)$$

where $i = 1, 2, \dots, n-1$,

n = the number of dynamic signs,

x_i = spacing between signs $i-1$ and i ,

k = spacing constant.

Equation 4.1 allows for $n-1$ equations to be created with n unknowns. In order to solve for all of the unknowns, another equation is needed. This equation is shown below:

$$d = x_1 + \dots + x_n \quad (4.2)$$

where:

d = distance from the static sign C to the furthest dynamic n th sign from the work zone (see Fig. 1.1),

n = number of dynamic signs.

Equations 4.1 and 4.2 constitute a system of n equations with n unknown variables x_i if the values d and k are known. The solution for x_1 is:

$$x_1 = d/(1 + k + k^2 + \dots + k^n). \quad (4.3)$$

The remaining x_i can be calculated using Equation 4.2. Thus, the location of each dynamic sign is determined by one parameter, k , called here forward *spacing constant*. The three parameters, number of boards n , distance for dynamic signs d , and spacing constant k , define configuration of the dynamic signs. All three are optimized for each simulated case. Please notice that $k < 1$ corresponds to the growing spacing between signs in the direction of traffic while $k > 1$ corresponds to the decreasing spacing between signs in the direction of traffic. The value $k = 1$ corresponds to the even spread of the dynamic signs.

The optimization of the parameters is accomplished by means of a sequential one-dimensional search. In the basic step of this method, one variable is changed with a constant interval while holding all the others constant until the value that minimizes the MOE is found. Then, the next variable from the list is taken to find its optimal value. The value found is then fixed while other variables are considered. The solution is found when all the variables are used.

4.5 VARIABLES' CONSTRAINTS

In order to control the relative scope of the parameter optimization, limits have been placed on some of the parameters.

The spacing of the static signs (Sign A through Sign C in Fig. 1.1) is set to 150 meters (500 feet). Additionally, the minimum spacing between any two of the remaining signs is also set to 150 meters. This minimum distance is justified by considering the time and distance necessary for a driver to respond to any one of the signs. The fact that dynamic sign 1 (Figure 1.1) is always activated, guarantees that the no passing zone will be a minimum of 600 meters (2,000 feet) from the point at which the discontinued lane is terminated. This minimum no passing zone was determined based on the comments from the Indiana State Police solicited by INDOT personnel during the summer of 1997. This minimum requirement provides law enforcement officials with a sufficient distance to stop violators before they enter the work zone.

The total number of signs in the system has a lower limit of three static signs immediately upstream of the work zone followed by two dynamic signs, for a total minimum of five signs. An upper limit is three static signs plus five dynamic signs.

4.6 SIMULATION RESULTS

First, a work zone without DO NOT PASS signs has been simulated for the four levels of traffic demand. The work zone's capacity was 1450 veh/h (as observed at Indiana locations). Simulation results are presented in Table 4.1. As could be expected, the travel times are growing with the growth of traffic demand. The travel times in the continuous lane are longer than in the discontinued lane, and this discrepancy increases when the traffic volume becomes heavier.

The average deceleration rate has a tendency to decrease with the increasing level of congestion (higher demand level). It seems that the most dangerous is a situation of traffic

transition for non-congested to congested conditions. This phase lasts a significant time where traffic demand is only slightly above the capacity. Vehicles approaching the work zone experience strong turbulence of speed. On the other hand, when the demand is significantly higher than capacity, the queue on the approach is well established and long. The speed variations, although still high, are lower than in the traffic transition from non-congested to congested states.

The simulation results obtained for the no control situation were used to calculate the change in the number of passing maneuvers and to evaluate the effectiveness of ILMS.

Table 4.1 Simulation Experiment (no control)

Volume (veh/h)	Outflow (veh)	System travel time (min)	Travel time in continuous lane (min)	Average deceleration rate (m/s^2)	Average no. of departures
1500	2635	6.39	6.84	0.307	150.3
1600	2703	8.17	9.95	0.287	224.1
1700	2707	9.95	12.91	0.232	291.2
1800	2707	11.62	14.81	0.191	334.1

Secondly, the same work zone was simulated in the presence of eight DO NOT PASS signs. A sequence of simulation was performed for each traffic demand level to determine the number of activated boards and to find the optimal spacing between boards. Travel time in the continuous lane and reduction in the number of passing maneuvers are the criteria used. Table 4.2 presents the performed four simulation experiments.

The first static board has fixed location in all simulations -- 150 m upstream of the taper. Two different spacing values were tested for the two other boards: 150 m and 300 m.

For each spacing of static boards, the uniform spacing of five dynamic boards has been evaluated starting at 150 m with 150-m increment. The search was terminated after the number of passing maneuvers increased. Finding the optimal spacing of dynamic boards

and the number of active boards completed the search for a given traffic demand. The optimal settings are in shaded lines in Table 4.2.

In all the optimal solutions, the travel times in the continuous lane and the reduction in the number of passing maneuvers reach minimum values. Although the two measures of a system's performance yield the same solution, the reduction in the number of passing maneuvers is more convenient for two reasons.

1. It is more sensitive to the system's settings than the travel time in the continuous lane.
2. It is a measure of the system's effectiveness since it relates the system's performance to the situation without DO NOT PASS boards.

The average deceleration rate behaves in a very consistent manner. It grows with the length of the segment controlled by the DO NOT PASS boards, but it never exceed the value obtained for the work zone without ILMS.

Figures 4.2 through 4.4 present the system's effectiveness in the function of the boards' spacing. Three important conclusions can be drawn from the results.

1. The system can increase the travel time in the continuous lane and the number of passing maneuvers when the system's setting is not optimal. This is particularly obvious for low traffic volume and short congested segment (1.5 km or shorter).
2. The ILMS set in optimal manner improves the work zone's performance compared to the situation without the ILMS.
3. The system's performance and effectiveness grows with the traffic demand. The benefit is obvious where the maximum congested segment exceeds 3 km.
4. Maximum number of dynamic boards should be used where the congested segment exceeds 3 km. The number of boards should be reduced if the maximum length of the congested segment is shorter than 3 km.

The spacing in the simulation experiments was assumed uniform for all cases. The effect of three different spacing patterns, represented by three different values of the spacing constant k , is shown in Table 4.3. The uniform spacing has turned out optimal. This finding simplifies the setting rules quite significantly.

The last setting that required investigation was the threshold detector occupancy. This setting in the simulation experiments was set at 30% -- a typical threshold for congested conditions used by freeway traffic centers. The sensitivity analysis performed for this setting produced results that indicated no significant effect of this value on the work zone's performance (see Table 4.4). The value of 30% remains the recommended value.

The simulation experiments have shown that ILMS is likely to improve users' perception of a work zone's performance. More importantly, ILMS is likely to reduce the number of crashes. The results of simulation experiments indicate the importance of the warrants of ILMS use and the rules of the system's setting.

4.7 PROPOSED RULES OF SYSTEM SETTING

The discussion of the simulation results has already pointed out several settings:

1. Spread static boards at a frequency of 300 meters,
2. Use the maximum number of dynamic boards,
3. Spread the dynamic boards in a uniform manner (equal spacing),
4. Use the threshold detector occupancy of 30%.

The results also have indicated that the system is quite sensitive to the spacing between dynamic boards, which is directly related to the distance between the farthest sign and the work zone. This value is called here forward "a deployment area." Following is a description of the procedure that returns the length of deployment area and the spacing between dynamic boards based on the maximum length of congested segment.

The length of the congestion segment is a function of the traffic density on the approach and the number of vehicles that could not pass the work zone due to the overflow. The latter value is sometimes called an overflow queue. The overflow queue reaches its maximum value at the end of the overflow period. It is equal to the number of vehicles that arrive during the overflow period minus the total capacity of the work zone entry during the overflow period. These vehicles accumulate on the approach at a certain density. The average traffic density on the approach to the work zone controlled by optimal IMLS has been estimated based on the results obtained from the simulation (see Table 4.5). This value has been determined as 71.6 veh/km (total value for two lanes).

The congestion segment is calculated as follows:

$$L_c = \frac{0.92 \cdot Q}{D_c - D} \quad (4.5)$$

where:

Q = maximum overflow queue, veh,

D_c = average traffic density along congested segment, $k_c = 71.6$ veh/km,

D = traffic density if there is no work zone, veh/km.

The adjustment 0.92 has been determined using the simulation results. The overflow vehicles Q add to the vehicles that would be on the approach in the absence of a work zone. That is why the traffic density D_c is reduced by the unaffected density D that would be observed if there was no work zone. Density D can be calculated using the HCM method for freeways. A simplified equation has been derived from the steady-state formula (Equation 2.27):

$$D = 46.3 - \sqrt{2150 - 0.742 \cdot V} \quad (4.6)$$

where V is traffic volume arriving during the last hour of the overflow period, (veh/h).

We formulate a hypothesis that an optimal deployment area depends on a maximum congested segment. Additional cases have been simulated to increase the simulated data

data to check this hypothesis. Four cases were different from the cases already considered by longer overflow periods (1.5 and 2 hours). The fifth case comprises a two-hour overflow period with the volume changing in half-hour intervals (1600, 1700, 1800, 1700 veh/h). The results are shown in Table 4.6. The deployment areas with the lowest number of passing maneuvers are in shaded lines.

The nine simulated cases with optimally set ILMS have been used to check the relationship between the length of congested segment and the optimal deployment area. Table 4.7 presents the maximum lengths of the congested segments calculated using the method earlier presented and the corresponding optimal deployment areas. There is indeed, a relationship between these to values as shown in Figure 4.5. The best fit has been obtained using the following equation:

$$d = 1.91 \cdot L_c^{0.56} \quad (4.4)$$

where:

d = deployment area (km),

L_c = congested segment (km).

The deployment area now can be easily calculated using the above sequence of equations. Part II of the report contains a complete set of rules of setting the ILMS.

Table 4.2 Simulation Experiments to Optimize the System (Threshold Occupancy = 30%, Min. Activation = 5 min.)

Volume (veh/h)	Max. # of active boards	Boards location (m)		Outflow (veh)	Travel time (min)		Average decel. rate (m/s ²)	No. of passing veh. (veh)	Change in passing (veh)
		Static	Dynamic		System	Continuous lane			
1500	8	150	150	2566	7.3	8.9	0.147	190	39.5
1500	6	150	300	2700	6.5	6.8	0.173	153	3.0
1500	4	150	600	2635	6.4	6.9	0.211	151	0.8
1500	4	150	900	2635	6.4	6.9	0.213	151	0.8
1500	4	150	1200	2635	6.4	6.9	0.213	151	0.7
1500	4	150	1500	2635	6.4	6.9	0.213	151	0.5
1500	8	300	150	2635	6.6	7.1	0.188	156	6.0
1500	5	300	300	2635	6.4	6.8	0.197	148	-1.9
1500	5	300	600	2635	6.4	6.8	0.211	149	-1.3
1500	5	300	900	2635	6.4	6.8	0.212	149	-1.4
1600	8	150	150	2462	10.8	14.3	0.081	294	70.1
1600	8	150	300	2566	9.5	12.1	0.100	259	34.6
1600	8	150	450	2610	8.9	11.0	0.114	240	16.0
1600	8	150	600	2645	8.6	10.4	0.124	229	5.0
1600	6	150	900	2705	8.2	9.9	0.150	224	-0.2
1600	5	150	1050	2703	8.2	10.3	0.163	232	7.5
1600	5	150	1200	2800	8.2	10.3	0.168	241	16.4
1600	5	150	1500	2703	8.2	10.4	0.175	234	10.4
1600	8	300	300	2637	8.7	10.7	0.118	236	11.5
1600	8	300	600	2635	8.3	9.9	0.130	217	-6.7
1600	6	300	900	2704	8.2	10.0	0.150	225	0.7

Table 4.2 Simulation Experiments to Optimize the System - continuation
(Threshold Occupancy = 30%, Min. Activation = 5 min.)

Volume (veh/h)	Max. # of active boards	Boards location (m)		Outflow (veh)	Travel time (min)		Average decel. rate (m/s ²)	No. of passing veh. (veh)	Change in passing (veh)
		Static	Dynamic		System	Continuous lane			
1700	8	150	300	2635	11.8	14.9	0.073	328	36.8
1700	8	150	600	2605	10.8	13.3	0.088	289	-2.5
1700	8	150	900	2647	10.3	13.0	0.100	288	-3.7
1700	8	150	1050	2663	10.2	12.4	0.105	276	-15.4
1700	7	150	1200	2655	10.1	12.9	0.110	284	-6.9
1700	6	150	1500	2707	10.0	13.1	0.119	297	5.4
1700	8	300	300	2596	10.8	13.7	0.085	297	6.0
1700	8	300	600	2660	10.4	12.8	0.094	283	-8.2
1700	8	300	750	2685	10.2	12.5	0.098	280	-11.0
1700	8	300	900	2702	10.0	12.2	0.101	275	-16.6
1700	8	300	1050	2709	10.0	12.2	0.104	276	-14.8
1700	7	300	1200	2706	10.0	12.5	0.106	282	-9.4
1800	8	150	150	2333	16.1	20.5	0.049	399	65.3
1800	8	150	300	2509	13.8	17.1	0.059	357	22.9
1800	8	150	450	2546	13.2	16.4	0.064	348	13.7
1800	8	150	600	2585	12.8	15.4	0.070	332	-1.9
1800	8	150	900	2625	12.3	14.7	0.078	321	-13.5
1800	8	150	1050	2640	12.1	14.2	0.080	313	-21.2
1800	8	150	1200	2655	11.9	14.5	0.082	321	-13.5
1800	7	150	1500	2686	11.8	14.6	0.092	327	-7.2
1800	8	300	300	2596	12.8	15.8	0.067	341	7.3
1800	8	300	600	2644	12.2	14.8	0.074	327	-7.5
1800	8	300	750	2705	11.7	14.4	0.079	323	-10.6
1800	8	300	900	2691	11.8	14.3	0.078	321	-12.8
1800	8	300	1050	2669	11.9	14.2	0.077	315	-18.9
1800	8	300	1200	2711	11.6	14.3	0.079	323	-11.4

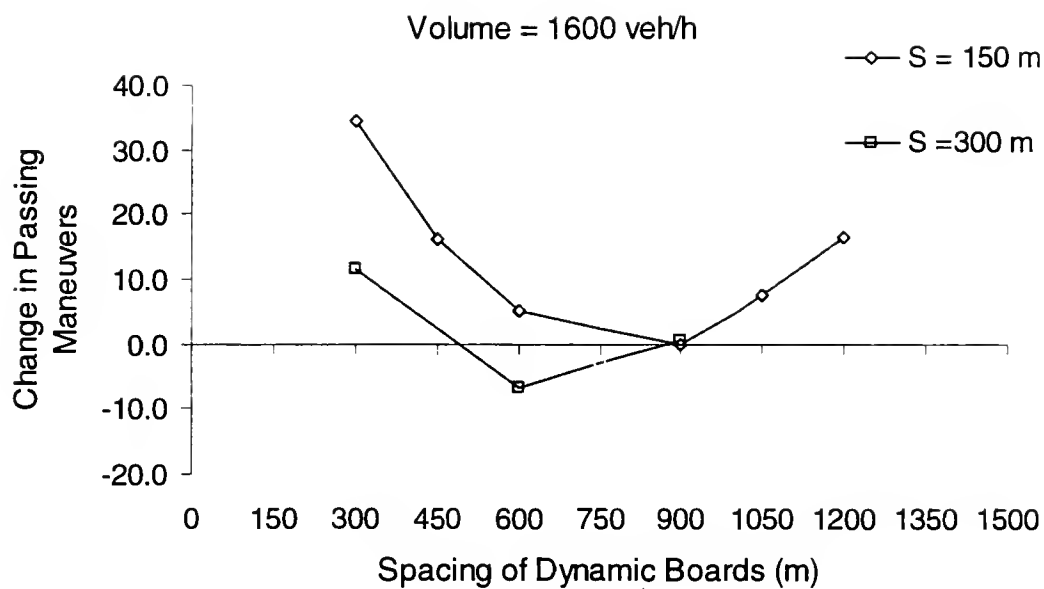


Figure 4.2 Results of the Simulation Experiments for $V = 1600$ veh/h

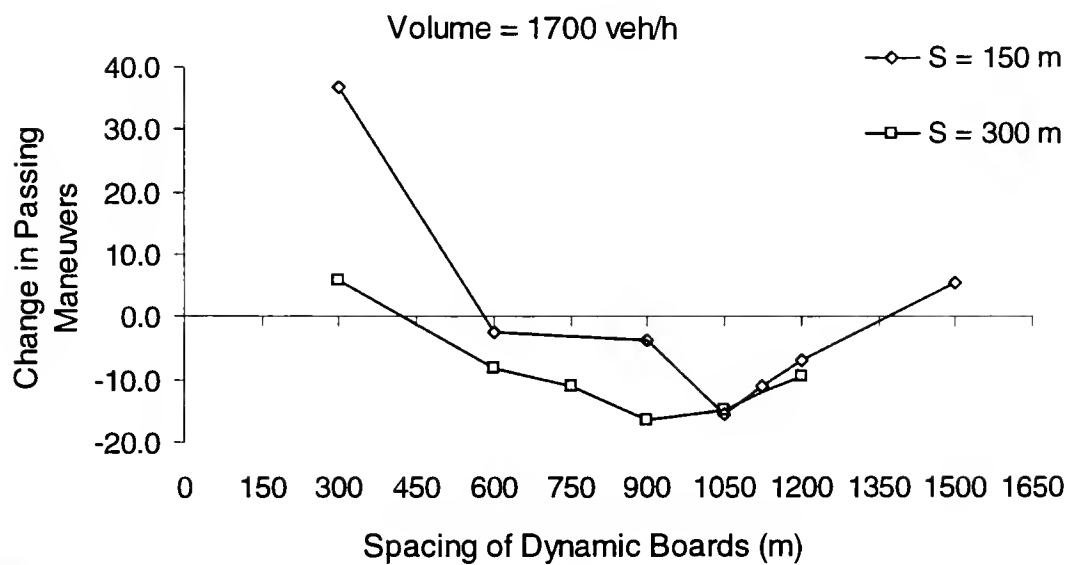


Figure 4.3 Results of the Simulation Experiments for $V = 1700$ veh/h

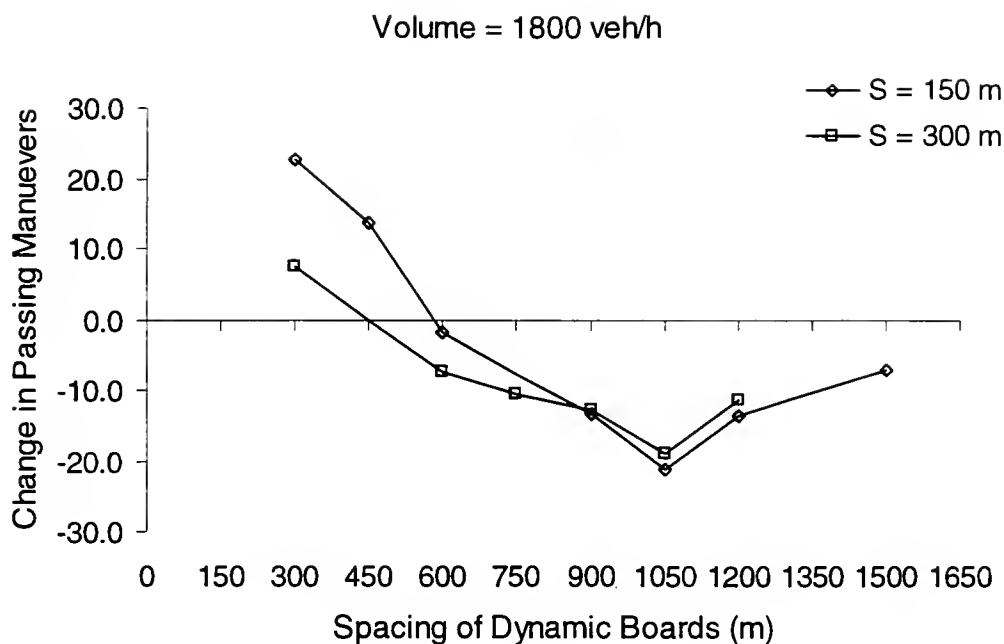


Figure 4.4 Results of the Simulation Experiments for $V = 1800$ veh/h

Table 4.3 Effect of Spacing Uniformity (spacing of static boards = 150, 450, 750m; total distance with dynamic boards = 3000 m; number of boards = 5).

Volume (veh)	Occupancy (%)	Spacing constant	Capacity (veh /hr)	Change in passing maneuvers (veh)	Deceleration rate (m/s^2)	Acceleration noise (m^2/s^2)	Travel time On continuous lane (min)
1700	30	0.8	1320	-25	0.1	108.5	9.5
1700	30	1.0	1320	-29	1.3	109.8	9.3
1700	30	1.2	1320	-25	0.1	108.5	9.5

Table 4.4 Effect of Threshold Occupancy (spacing of static boards = 300 m)

Volume (veh/h)	Spacing of dynamic boards (m)	Occupancy (%)	Travel time (min)		Average deceleration rate (m/s^2)
			Continuous lane	System	
1600	750	25	9.5	8.2	0.137
		30	9.5	8.2	0.139
		35	9.4	8.2	0.138
1700	900	25	12.2	10.0	0.101
		30	12.2	10.0	0.101
		35	12.2	10.0	0.102
1800	1050	25	14.2	11.7	0.079
		30	14.2	11.9	0.077
		35	14.3	11.7	0.079

Table 4.5 Average Traffic Densities on Approaches to Work Zones with Optimal ILMS

Case	Overflow volume (veh/h)	Max. congested segment (m)	Segment occupancy (veh)			Traffic density		
						Contin. lane	Discont. lane	Total density (veh/km)
						(veh/km/lane)		
1	1500	1200	69.0	24.0	93.0	57.5	20.0	77.5
2	1600	3050	169.5	47.2	216.7	55.6	15.5	71.0
3	1700	5050	273.0	70.1	343.1	54.1	13.9	67.9
4	1800	6600	353.0	106.6	459.6	53.5	16.2	69.6
			Average densities (veh/km)			55.2	16.4	71.6

Table 4.6 Additional Simulation Experiments

Case	Initial traffic		Overflow traffic		Final traffic		Boards location (m)		Outflow (veh)	Travel time (min.)		Decel. rate (m/s ²)	No. of passing veh.	No. of passages above min.
	Volume (veh/h)	Period (h)	Volume (veh/h)	Period (h)	Volume (veh/h)	Period (h)	Static	Dynamic		Time on Cont. lane	System time			
5	1400	0.5	1700	1.5	800	1	300	850	4802	13.36	11.23	0.068	356.4	14.4
	1400	0.5	1700	1.5	800	1	300	1000	4802	12.82	11.06	0.070	342.0	0.0
	1400	0.5	1700	1.5	800	1	300	1150	4802	13.14	11.13	0.072	350.4	8.4
6	1400	0.5	1700	2	800	1.5	300	950	5230	21.73	17.12	0.054	473.5	1.9
	1400	0.5	1700	2	800	1.5	300	1100	5230	21.64	17.04	0.056	471.6	0.0
	1400	0.5	1700	2	800	1.5	300	1250	5230	21.79	17.24	0.057	474.8	3.2
	1400	0.5	1800	1.5	800	1	300	1150	4035	25.35	19.85	0.050	568.3	5.1
7	1400	0.5	1800	1.5	800	1	300	1300	4027	25.30	19.92	0.050	566.1	2.9
	1400	0.5	1800	1.5	800	1	300	1450	4024	25.19	19.95	0.050	563.1	0.0
	1400	0.5	1800	1.5	800	1	300	1600	4023	25.30	19.96	0.051	565.4	2.3
8	1400	0.5	1800	2	800	1.5	300	1300	5409	31.34	23.66	0.040	706.3	1.7
	1400	0.5	1800	2	800	1.5	300	1450	5408	31.27	23.73	0.040	704.6	0.0
	1400	0.5	1800	2	800	1.5	300	1600	5408	31.37	23.75	0.041	706.8	2.3
9	Inflow volumes for 0.5-hour intervals (veh/h): 1400, 1600, 1700, 1800, 1700, 800, 800, 800.						300	800	5230	21.45	16.75	0.059	467.4	14.1
							300	950	5230	20.80	16.52	0.060	453.3	0.0
							300	1100	5230	20.92	16.52	0.061	455.9	2.6
							300	1250	5230	21.00	16.66	0.061	457.3	4.0

Table 4.7 Calculated Maximum Congested Segments versus Optimal Deployment Areas

Case	Total capacity TC (veh)	Maximum overflow queue Q (veh)	Last hour volume q (veh/h)	Unaffected traffic density k_o (veh/h)	Calculated maximum congested zone L_c (km)	Simulated optimal deployment zone d (km)
1	1420	80	1500	14.10	1.28	1.35
2	1420	180	1600	15.27	2.94	3.75
3	1410	290	1700	16.49	4.84	5.25
4	1410	390	1800	17.76	6.66	6.00
5	2130	420	1700	16.49	7.01	6.30
6	2840	560	1700	16.49	9.35	7.28
7	2130	570	1800	17.76	9.74	7.43
8	2840	760	1800	17.76	12.99	8.58
9	2840	560	1700	16.49	9.35	7.28

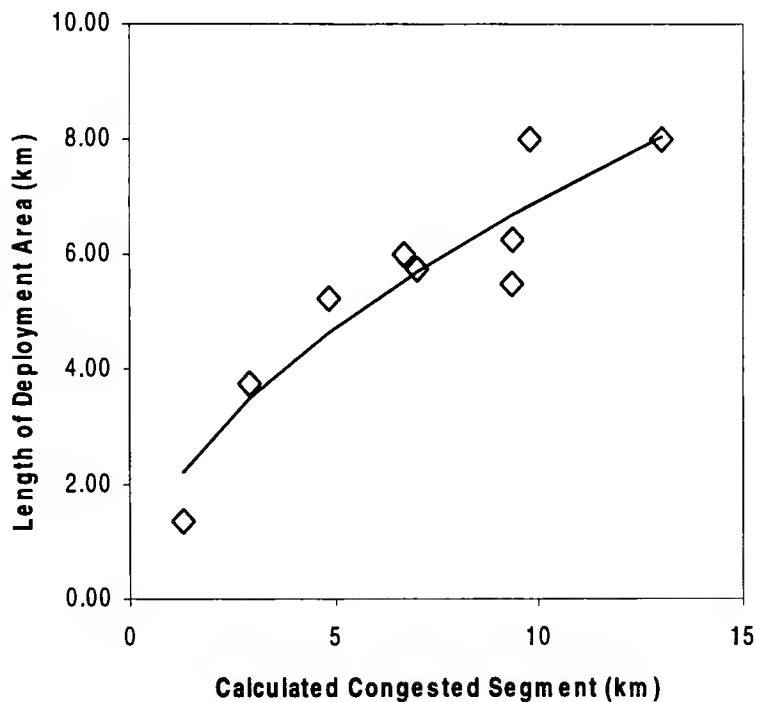


Figure 4.5 Maximum Congested Segment versus Deployment Area.

